

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:14:13 ; Search time 29.79 Seconds  
(without alignments)  
894,854 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 1304  
Sequence: 1 NGBRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNIRDEL 240

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802:\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1304	100.0	258	6 AAP50190	Sequence encoded b
2	1298	99.5	258	6 AAP50191	Sequence encoded b
3	1298	99.5	382	22 AAU00506	E. coli heat-labli
4	1293	99.2	240	19 AAW65074	E. coli LT-A prote
5	1288	98.8	240	19 AAW65075	E. coli LT-A mutan
6	1287	98.7	259	21 AAY96646	Plant-optimized E.
7	1283	98.4	259	21 AAY96647	Synthetic E. coli
8	1282	98.3	259	21 AAY96649	Plant-optimized E.
9	1280	98.2	259	21 AAY96648	Plant-optimized E.
10	1280	98.2	380	22 AAU00507	E. coli heat-labli
11	1276	97.9	259	21 AAY96650	Plant-optimized E.

12	1275	97.8	259	21	AA95651	Plant-optimized E.
13	1190.5	91.3	237	20	AAW67772	E. coli heat labli
14	1183	90.7	236	14	AAAR38728	E. coli heat labile
15	1183	90.7	254	22	AAU14105	Peptide sequence f
16	1179	90.4	236	14	AAAR44016	"Lys-63" E.coli he
17	1179	90.4	236	14	AAAR44024	"Glu-114" E.coli h
18	1179	90.4	236	14	AAAR44025	"Lys-114" E.coli h
19	1178	90.3	236	14	AAAR44018	"Tyr-37" E.coli h
20	1178	90.3	236	14	AAAR38732	"Tyr-53" E.coli he
21	1177	90.3	236	14	AAAR44017	"Lys-97" E.coli he
22	1177	90.3	236	14	AAAR38731	"Glu-53" E.coli he
23	1176	90.2	236	14	AAAR38730	"Asp-53" E.coli he
24	1175	90.1	236	14	AAAR44019	"Glu-107" E.coli h
25	1175	90.1	236	14	AAAR44023	"Lys-106" E.coli h
26	1174	90.0	236	14	AAAR44020	"Ser-104" E.coli h
27	1174	90.0	236	14	AAAR44022	"Ser-104" E.coli h
28	1173	90.0	236	14	AAAR44021	"Asp-104" E.coli h
29	1088	83.4	240	14	AAAR38729	Cholera toxin subu
30	1088	83.4	240	19	AAW80807	Amino acid sequenc
31	1088	83.4	258	21	AA95653	Plant-optimized V.
32	1088	83.4	258	22	AAAG65991	Cholera toxin A su
33	1087	83.4	258	21	AA95655	Plant-optimized mu
34	1085	83.2	240	20	AAW67773	Cholera toxin subu
35	1085	83.2	258	12	AAAR13117	Cholera toxin Al f
36	1084	83.1	240	14	AAAR44027	Lys-63 cholera tox
37	1084	83.1	240	14	AAAR44033	Glu-114 cholera to
38	1084	83.1	258	21	AA95654	Plant-optimized mu
39	1083	83.1	240	14	AAAR44031	Ser-110 cholera to
40	1082	83.0	240	14	AAAR44028	Lys-97 cholera tox
41	1082	83.0	240	14	AAAR44032	Ala-112 cholera to
42	1081	82.9	240	14	AAAR44030	Asn-107 cholera to
43	1081	82.9	240	14	AAAR44026	Asp-53 cholera tox
44	1081	82.9	258	21	AA95656	Plant-optimized mu
45	1080	82.8	240	14	AAAR44029	Ser-106 cholera to

ALIGNMENTS

RESULT	1
AAAP50190	
ID	AAAP50190 standard; Protein; 258 AA.
XX	AC
XX	AAAP50190;
XX	AC
XX	30-OCT-1991 (first entry)
XX	DE
XX	Sequence encoded by the pig scours heat labile toxin (LT) LTA gene.
XX	XX
XX	Pig scours vaccine; toxin; diarrhoea.
XX	OS
XX	E. coli NCIB 11932.
XX	XX
XX	EP145486-A.
XX	19-JUN-1985.
XX	XX
XX	12-DEC-1984; 84EP-0308620.
XX	XX
XX	12-DEC-1983; 83GB-0033131.
XX	(GLAX ) GLAXO GROUP LTD.
XX	XX
XX	Hayes MV, Harford S, Ross GW;
XX	WPI; 1985-148358/25.
XX	DR N-PSDB; AAN50205.
XX	XX
XX	New toxoid as inactivated form of toxin for use in vaccines - is
XX	obtd. from organism transformed by gene
XX	Disclosure; Fig 1; 61pp; English.
XX	XX

CC AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of  
 CC the site directed mutant SDM1 (see AAN50206) is inactive. The  
 CC inventors claim a vaccine prepn. active against pig scours which  
 CC contains an inactivated LTA component, together with additional K88  
 CC antigens opt. with whole cells comprising the antigens or contg. the  
 CC inactivated LTA.

SQ Sequence 258 AA;

Query Match 100.0%; Score 1304; DB 6; Length 258;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-127; Indels 0; Gaps 0;  
 Matches 240; Conservative 0; Mismatches 0;

QY 1 NGDRLRADSRPPDEIKRSGGLMPRGHNEYFDRTGMNINLYDHARGTGTGFVRYDDGYV 60  
 Db 19 ngdrlradsrppdeikrsgglmprghneyfdrtgmnlndhargtgtgfvryddgyv 78  
 QY 61 STSLRSALHLAGOSILSGYSTIYIVATAPNMFVNNDVGLGYSPHPYEQEVSALGGIP 120  
 Db 79 stslrsalhlagsilsgystiylviatapnmfnvndvlgvysphpyeqevsalggip 138  
 QY 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGFPDPHQAWREEPWI 180  
 Db 139 ysqiygwyrvnfgviderlhrnreyrdyrrnlnlapiadegyrlagfpdphqawreepwi 198  
 QY 181 HPAQCGNSSRTITGDTCEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240  
 Db 199 hpaqcgcnssrtitgdtcneetqnlstilyreyqskvkrqifsdysqvsevdynrirdel 258

#### RESULT 2

ID AAP50191 standard; Protein; 258 AA.

AC AAP50191;

DT 30-OCT-1991 (first entry)

DE Sequence encoded by the pig scours heat labile toxin (LT) LTA gene  
 DE of the site directed mutant SDM1.

KW Pig scours vaccine; toxin; diarrhoea.

OS E.coli NCIB 11932.

XX Key Location/Qualifiers

FT Misc-difference 79 /note= "Ser in native SQ"

XX EP145486-A.

XX 19-JUN-1985.

XX 12-DEC-1984; 84EP-0308620.

XX 12-DEC-1983; 83GB-0033131.

XX (GLAX ) GLAXO GROUP LTD.

XX Hayes MV, Harford S, Ross GW;

XX WPI; 1985-148358/25.

XX N-PSDB; AAN50206.

XX New toxoid as inactivated form of toxin for use in vaccines - is  
 XX obtd. from organism transformed by gene

PS Example; Fig 2; 61pp; English.

XX AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of  
 CC the site directed mutant SDM1 (see AAN50206) is inactive. The  
 CC inventors claim a vaccine prepn. active against pig scours which

CC contains an inactivated LTA component, together with additional K88  
 CC antigens opt. with whole cells comprising the antigens or contg. the  
 CC inactivated LTA.

SQ Sequence 258 AA;

Query Match 99.5%; Score 1298; DB 6; Length 258;  
 Best Local Similarity 99.6%; Pred. No. 1.2e-126;  
 Matches 239; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGDRLRADSRPPDEIKRSGGLMPRGHNEYFDRTGMNINLYDHARGTGTGFVRYDDGYV 60  
 Db 19 ngdrlradsrppdeikrsgglmprghneyfdrtgmnlndhargtgtgfvryddgyv 78  
 QY 61 STSLRSALHLAGOSILSGYSTIYIVATAPNMFVNNDVGLGYSPHPYEQEVSALGGIP 120  
 Db 79 stslrsalhlagsilsgystiylviatapnmfnvndvlgvysphpyeqevsalggip 138  
 QY 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGFPDPHQAWREEPWI 180  
 Db 139 ysqiygwyrvnfgviderlhrnreyrdyrrnlnlapiadegyrlagfpdphqawreepwi 198  
 QY 181 HPAQCGNSSRTITGDTCEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240  
 Db 199 hpaqcgcnssrtitgdtcneetqnlstilyreyqskvkrqifsdysqvsevdynrirdel 258

#### RESULT 3

AAU00506 standard; Protein; 382 AA.

XX AAU00506;

DT 29-AUG-2001 (first entry)

DE E. coli heat-labile enterotoxin (LT) mutant LTS63Y.

KW Heat-labile enterotoxin; LT; LTS63Y; Ltdell10/112; mutant;

KW detoxified and immunologically active protein; ADP-ribosylation; Gs;  
 KW endotoxin; diarrhoea; mutein.

XX Escherichia coli strain K88ac.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc\_feature 7 /note= "Important residue for enzymatic activity"

FT Misc\_feature 44 /note= "Important residue for enzymatic activity"

FT Region 58..72 /note= "Forms the NAD-binding site"

FT Misc\_feature 61 /note= "Important residue for enzymatic activity"

FT Misc-difference 63 /note= "Substitution of wild type Ser to Tyr"

FT Misc\_feature 110 /note= "Important residue for enzymatic activity"

FT Misc\_feature 112 /note= "Important residue for enzymatic activity"

FT Misc-difference 259 /note= "Encoded by TG"

XX WO200119998-A1.

XX 22-MAR-2001.

XX 15-SEP-1999; 99WO-KR00555.

XX 15-SEP-1999; 99WO-KR00555.

XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.

XX

PI	Park EJ, Kim JS, Chang J, Yum J, Chung S;
XX	WPI: 2001-281524/29.
DR	N-PSDB; AAS01505.
XX	
PT	New detoxified mutants of <i>Escherichia coli</i> heat-labile enterotoxin
PT	useful as vaccine for preventing and treating diarrhoea, and as adjuvant
PT	for antibody production
PS	
XX	Claim 2; Page 39-41; 48pp; English.
XX	
CC	The present sequence represents <i>Escherichia coli</i> heat-labile
CC	enterotoxin (LT) mutant LT563Y. LT563Y and LTdel110/112 (AAU00507)
CC	are two novel detoxified and immunologically active proteins (LT
CC	mutants) derived by site-directed mutagenesis of the A1 subunit of wild
CC	type LT. The substitution of Ser to Tyr at position 63 in LT563Y blocks
CC	NAD-binding. Deletion of Glu residues at positions 110 and 112 in
CC	LTdel110/112 eliminate the enzymatic activity of LT. The A1 subunit of
CC	wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that
CC	regulates cAMP levels. The resulting increase in cAMP is the cause of
CC	diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities
CC	of mutant heat-labile endotoxins LT563Y and LTdel110/112 were tested.
CC	Groups of mice were immunised with LT5623Y or LTdel110/112. The control
CC	groups received phosphate buffered saline (PBS) alone. The serum and
CC	faecal antibody titres to LT were determined. The results showed that
CC	mice immunised with LT563Y or LTdel110/112 contained high and
CC	comparable level of anti-LT antibodies in sera and faecal extracts
CC	compared with those immunised with wild-type LT. The LT mutants are
CC	useful as a vaccine for preventing and treating diarrhoea and as an
CC	adjuvant for antibody production.
XX	
XX	Sequence 382 AA;
XX	
QY	Query Match 99.5%; Score 1298; DB 22; Length 382;
DB	Best Local Similarity 99.6%; Pred. No. 2.2e-126;
DB	Matches 239; Conservative 0; Mismatches 1; Indels 0; Gaps
QY	1 NGRLYRADSRPDEIKRSGGLMPRGHNHYFDRGTQMNINLYDHARGTGTGFVYDDGYV 60
DB	19 ngdriyradsrpdeldkrsggimprghneyfdrgtgmnlndydhargtqgtgfvyddgyv 78
QY	61 SFLSLRSALHAGOSTLSCYSTSYIYVIATAPNMFVNDVLGYSPHPYQEVSAIIGGIP 120
DB	79 stylsrlrsahlagqslsgystyiyviatapnmfndvlgvysphpyeqevsaigglp 138
QY	121 YSQIYCWYRVNFGVIDERLHNRREYDRYRNLNIAPAEDGYRLAGFPDQAWREPWI 180
DB	139 ysqiygwyrvvnfgvidelrhnrreaydrdryrnlndyapadgyrlagfpdqawreepwi 198
QY	181 HHAPOGCGNSRTTIGDTCNEETQNLSTLYLREYQSKVRKQIFSDYQSEVDIYNRDEL 240
DB	199 hhapogcgnsrttigtcdneetqnlstlylreyqskvrqifdsdyqsevdinyrirdel 258
RESULT	4
AAW65074	
ID	AAW65074 standard; protein; 240 AA.
XX	
AC	AAW65074;
XX	
DT	11-SEP-1998 (first entry)
XX	
DE	E. coli LT-A protein fragment.
XX	
KW	Heat labile toxin subunit A; LT-A; mutant; immunogen; detoxification;
KW	carrier; adjuvant; prevention; treatment; disease; diarrhoea; vaccine;
KW	infection; enterotoxic.
XX	
OS	<i>Escherichia coli</i> .
XX	
FT	Key Location/Qualifiers
PH	Protein 1..241

```

/note= "partial sequence"
WO9818928-A1.
07-MAY-1998.
30-OCT-1997; 97WO-IB01440.
31-OCT-1996; 96GB-0022660.
(CHIR-) CHIRON SPA.
Giuliani MM, Pizza M, Rappuoli R;
WPI; 1998-272223/24.
Mutated Escherichia coli heat labile toxin subunit A - is immunogenic and detoxified relative to wild-type, useful e.g. in vaccines against E. coli enterotoxigenic strains and as an adjuvant
Disclosure; Page -: 67pp; English.
This sequence represents a fragment of an Escherichia coli heat labile toxin subunit A (LT-A). This protein is used in a method resulting in a mutant LT-A protein which has the wild type Ala residue at position 72 changed to an Arg residue resulting in a toxin which retains its immunogenicity but is detoxified. Detoxification is defined in the specification as a reduction in toxicity relative to wild-type toxin, such that any residual toxicity is sufficiently low to allow use as an effective immunogenic composition in humans without significant side effects. The protein can be combined with an acceptable carrier in immunogenic compositions, optionally comprising an adjuvant and/or a second immunogenic antigen. Such compositions can be administered to prevent/treat disease in a subject e.g. traveller's diarrhoea in humans. The protein or compositions are especially administered as vaccines useful to prevent or treat infections by enterotoxigenic strain of E. coli in mammals (especially humans).
Sequence 240 AA;
Query Match          99.2%;   Score 1293;   DB 19;   Length 240;
Best Local Similarity 99.2%;   Pred. No. 3.7e-126;
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy      1 NGDRLYRADSRPPDEIKRSGGLMPGRGNEYFDRTGTQMNLNYLDHARTQTGFVRYDDGYV 60
Db      1 ngdrlyradsrppdeikrs gglmpgrghneyfdrtgtqmnl nyldhargtqtgfvyddgyv 60
Qy     61 STSLSLRAHLAQSILSGYSTIYYIVTATAPNPNVNVDLGVISPHYPDEVSAALGGIP 120
Db     61 stslslrahlagqsilsgyst iyyivtatapn pnvndl gvvisphypdevsalggip 120
Qy    121 YSQLYGWVRVFNGVIDERLHNREYDRDYRNLANTPADECGYRLAGPDPDHQAWREPWI 180
Db    121 ysqlygwvrvfngvider lhnre ydrdy rnylnatpadec gyrlagpdpdhqawrepwi 180
Qy    181 HHAPOGCNSRSRTITGDTCNSETONLSITYLREYSQSKVKRFISDYQSEVDIYNRIDEL 240
Db    181 hhapogcnsrtsitgd tcncsetqnls itylrey sqskvkrf ifsy qsevd ynri del 240
RESULT           5
ID       AAW65075
XX       ID AAW65075 standard; protein; 240 AA.
AC       XX AAW65075;
XX       AC AAW65075;
DT       DT 11-SEP-1993 (first entry)
XX       DE E. coli LT-A mutant A72R protein fragment.
XX       WW Heat labile toxin subunit A; LT-A; mutant; immunogen; detoxification;
```



Db 80 stslslrsahlagsgllsgystylyviatapnmfnvndvlgvysphpyeqevsalgip 139  
Qy 121 YSQIYGWYRVNFGVTDRLHRNREYDRYRNLTAPAEDGYRLAGFPDHOAWREEPWI 180  
Db 140 ysqlygywrvnfgvtdrlhrnreydrfyrnlnlapeadgyrlagfpdhdqawreepwi 199  
Qy 181 HHAPQGCNSRRTTGDTCNEETQNLSTIYLREYOSKVKRQIFSDYQSEVDIYNRIDEL 240  
Db 200 hhapqgcgdsrtrttdtcneetqnlstlyrkqygvkrqifsdysgsevdlynrinel 259

RESULT 7  
ID AAY96647 standard; Protein: 259 AA.  
XX AAY96647;  
AC AAY96647;  
DT 26-SEP-2000 (first entry)  
XX Synthetic E. coli LT-A K63 mutant.  
DE Heat-labile toxin; LT-A; LT-B; mutein; transgenic plant; vaccine; oral;  
KW adjuvant; anti-bacterial; S63k.  
XX Escherichia coli.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Peptide 1..19  
FT /label= signal\_peptide  
FT /note= "20"  
FT Protein 20..259  
FT /label= mature\_protein  
FT Misc-difference 82  
FT /label= S63K  
FT /note= "Wild-type serine is replaced by lysine"  
XX  
XX WO200037609-A2.  
XX 29-JUN-2000.  
XX 22-DEC-1999; 99WO-US30747.  
XX 22-DEC-1998; 98US-0113507.  
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.  
XX (MASO/) MASON H S.  
XX (ARNT/) ARNTZEN C J.  
XX Mason HS, Arntzen CJ;  
XX WPI: 2000-442653/38.  
XX N-PSDB: AAA51147.  
XX New polynucleotides encoding LT-A or CT-A polypeptides for the  
XX transformation of plant cells, useful in immunogenic compositions to  
XX elicit immune responses in animals  
XX Example 2; Page -: 103pp; English.

This is mutant S63K Escherichia coli heat-labile toxin (LT) A subunit (LT-A). The wild-type serine was replaced with lysine at residue 63 of the mature protein, which was caused by a codon change of TCC to AAG in the coding sequence. The sequence contains plant-preferred codons and eliminates sequence motifs associated with spurious mRNA processing. A single codon insertion (GTG encoding valine) was made to accommodate the creation of a NcoI restriction site around the initiator methionine codon. Novel polynucleotides encode a mutant LT-A polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A) polypeptide, which have reduced enzyme activity as compared to the wild-type LT-A or CT-A polypeptide and where at least one of the codons is altered to a plant preferred codon. The polynucleotide further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B subunit (CT-B). The

CC polynucleotides are useful for the transformation of plant cells for the  
CC production of transgenic plants to produce edible vaccines, especially  
CC oral vaccines in transgenic plants for the prophylactic or therapeutic  
CC treatment against E. coli or V. cholerae. The mutant polypeptides are  
CC also useful as adjuvants.  
CC NB: This sequence does not appear in the specification, it was made from  
CC the wild type sequence shown in AAY96647, which appears in Figure 1.  
XX  
SQ Sequence 259 AA;

Query Match 98.4%; Score 1283; DB 21; Length 259;  
Best Local Similarity 97.9%; Pred. No. 4.5e-125;  
Matches 235; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRTQMNNINLYDHARGTGTGFVRYDDGYV 60  
Db 20 ngdklyradsrppdeikrsgglmprghneyfdrtqtmnninlydhargtgtgfvryddgyv 79  
Qy 61 STSLSLRSALHAGOSILSGYSTYIYVIATAPNMFNVDVLGVSYPHPEQEVSAALGIP 120  
Db 80 stklslrsahlagsgllsgystylyviatapnmfnvndvlgvysphpyeqevsalgip 139  
Qy 121 YSQIYGWYRVNFGVTDRLHRNREYDRYRNLTAPAEDGYRLAGFPDHOAWREEPWI 180  
Db 140 ysqlygywrvnfgvtdrlhrnreydrfyrnlnlapeadgyrlagfpdhdqawreepwi 199  
Qy 181 HHAPQGCNSRRTTGDTCNEETQNLSTIYLREYOSKVKRQIFSDYQSEVDIYNRIDEL 240  
Db 200 hhapqgcgdsrtrttdtcneetqnlstlyrkqygvkrqifsdysgsevdlynrinel 259

RESULT 8  
ID AAY96649 standard; Protein: 259 AA.  
XX AAY96649;  
AC AAY96649;  
DT 26-SEP-2000 (first entry)  
XX Plant-optimized E. coli LT-A R72 mutant.  
DE Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;  
KW adjuvant; anti-bacterial; A72R.  
XX Escherichia coli.  
OS Synthetic.  
XX Key Location/Qualifiers  
FH Peptide 1..19  
FT /label= signal\_peptide  
FT Protein 20..259  
FT /label= mature\_protein  
FT Misc-difference 91  
FT /label= A72R  
FT /note= "Wild type alanine is replaced by arginine"  
XX WO200037609-A2.  
XX 29-JUN-2000.  
XX 22-DEC-1999; 99WO-US30747.  
XX 22-DEC-1998; 98US-0113507.  
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.  
XX (MASO/) MASON H S.  
XX (ARNT/) ARNTZEN C J.  
XX Mason HS, Arntzen CJ;  
XX WPI: 2000-442653/38.  
XX N-PSDB: AAA51545.  
DR

XX New polynucleotides encoding LT-A or CT-A polypeptides for the  
PT transformation of plant cells, useful in immunogenic compositions to  
PT elicit immune responses in animals  
XX  
PS Example 4; Page -; 103pp; English.  
XX  
CC This is mutant A72R Escherichia coli heat-labile toxin (LT) A subunit  
CC (LT-A). The wild-type alanine was replaced with arginine at residue 72  
CC of the mature protein, which was caused by a nucleotide change of GC to  
CC AG at position 273-274 in the coding sequence. The sequence contains  
CC plant-preferred codons and eliminates sequence motifs associated with  
CC spurious mRNA processing. A single codon insertion (GTG encoding valine)  
CC was made to accommodate the creation of a NcoI restriction site around  
CC the initiator methionine codon. Novel polynucleotides encode a mutant  
CC LT-A polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A  
CC subunit (CT-A) polypeptide, which have reduced enzyme activity as  
CC compared to the wild-type LT-A or CT-A polypeptide and where at least one  
CC of the codons is altered to a plant preferred codon. The polynucleotide  
CC further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or  
CC a CT B subunit (CT-B). The polynucleotides are useful for the  
CC transformation of plant cells for the production of transgenic plants to  
CC produce edible vaccines, especially oral vaccines in transgenic plants  
CC for the prophylactic or therapeutic treatment against E. coli or V.  
CC cholerae. The mutant polypeptides are also useful as adjuvants.  
CC Note: This sequence does not appear in the specification. It was  
CC constructed from the wild type LT-A shown in AAY96646 which is given  
CC in Figure 1 of the specification.  
XX  
SQ Sequence 259 AA;

Query Match 98.3%; Score 1282; DB 21; Length 259;  
Best Local Similarity 97.9%; Pred. No. 5.7e-125;  
Matches 235; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
QY : 1 NGDRLYRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTGTGFVRYDDGYV 60  
Db 20 ngdklyradsrppdeikrsgglmpgrhneyfdrgtqmninlydhargtgtgfvryddgyv 79  
QY 61 STSLSRSAHLAQSIILSGYSTYIYVVIATAPNMNNDVLGVYSPHPYEQEVSALGGIP 120  
Db 80 stslsrslrshahlagqslilsgystyiyviatapnmfnvndvlgvysphpyeqevsalggip 139  
QY 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLAPAEQGYRLAGFPDPHQAWREEPWI 180  
Db 140 ysqiygywrvnfgviderlhrnreaydryrnlnlapeaqgyrlagfpdpdqawreepwi 199  
QY 181 HHAPQCGNSRRTITGTCNEETQNLSTIYLRQYQSKVRQIFSDYQSEVDIYNRIDEL 240  
Db 200 hhapqcgdsrsrtitgdtcneetqnlstlylrkyqskvrqifsdysqsevdynrinel 259

RESULT 9  
AAY96648  
ID AAY96648 standard; Protein; 259 AA.  
XX  
AC AAY96648;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Plant-optimized E. coli LT-A G192 mutant.  
XX  
DE Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;  
KW adjuvant; anti-bacterial; R192G.  
XX  
OS Escherichia coli.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
FT /label= signal\_peptide  
FT Protein 20..259

FT Misc-difference 211 /label= mature\_protein  
FT FT /label= R192G  
FT FT /note= "Wild-type arginine is replaced by glycine"  
XX  
PN WO200037609-A2.  
XX  
XX 29-JUN-2000.  
PD  
XX  
XX 22-DEC-1999; 99WO-US30747.  
PF  
XX  
XX 22-DEC-1998; 98US-0113507.  
PR  
XX  
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.  
PA (MASO/) MASON H S.  
PA (ARNT/) ARNTZEN C J.  
XX  
XX Mason HS, Arntzen CJ;  
PI  
XX  
XX WPI; 2000-442653/38.  
DR  
XX  
XX N-PSDB; AAA51544.  
DR  
XX  
XX New polynucleotides encoding LT-A or CT-A polypeptides for the  
PT transformation of plant cells, useful in immunogenic compositions to  
PT elicit immune responses in animals  
XX  
XX Example 3; Page -; 103pp; English.  
XX  
CC This is mutant R192G Escherichia coli heat-labile toxin (LT) A subunit  
CC (LT-A). The wild-type arginine was replaced with glycine at residue 192  
CC of the mature protein, which was caused by a codon change of TCC to AAG  
CC in the coding sequence. The sequence contains plant-preferred codons and  
CC eliminates sequence motifs associated with spurious mRNA processing. A  
CC single codon insertion (GTG encoding valine) was made to accommodate the  
CC creation of a NcoI restriction site around the initiator methionine  
CC codon. Novel polynucleotides encode a mutant LT-A polypeptide or a  
CC mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A) polypeptide,  
CC which have reduced enzyme activity as compared to the wild-type LT-A or  
CC CT-A polypeptide and where at least one of the codons is altered to a  
CC plant preferred codon. The polynucleotide further comprises a nucleic  
CC acid sequence encoding LT B subunit (LT-B) or a CT B subunit (CT-B). The  
CC polynucleotides are useful for the transformation of plant cells for the  
CC production of transgenic plants to produce edible vaccines, especially  
CC oral vaccines in transgenic plants for the prophylactic or therapeutic  
CC treatment against E. coli or V. cholerae. The mutant polypeptides are  
CC also useful as adjuvants.  
CC NB: This sequence does not appear in the specification, it was made from  
CC the wild type sequence shown in AAY96647, which appears in Figure 1.  
XX  
SQ Sequence 259 AA;

Query Match 98.2%; Score 1280; DB 21; Length 259;  
Best Local Similarity 97.9%; Pred. No. 9.3e-125;  
Matches 235; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGDRLYRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTGTGFVRYDDGYV 60  
Db 20 ngdklyradsrppdeikrsgglmpgrhneyfdrgtqmninlydhargtgtgfvryddgyv 79  
QY 61 STSLSRSAHLAQSIILSGYSTYIYVVIATAPNMNNDVLGVYSPHPYEQEVSALGGIP 120  
Db 80 stslsrslrshahlagqslilsgystyiyviatapnmfnvndvlgvysphpyeqevsalggip 139  
QY 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLAPAEQGYRLAGFPDPHQAWREEPWI 180  
Db 140 ysqiygywrvnfgviderlhrnreaydryrnlnlapeaqgyrlagfpdpdqawreepwi 199  
QY 181 HHAPQCGNSRRTITGTCNEETQNLSTIYLRQYQSKVRQIFSDYQSEVDIYNRIDEL 240  
Db 200 hhapqcgdsrsrtitgdtcneetqnlstlylrkyqskvrqifsdysqsevdynrinel 259

RESULT 10  
AAU00507  
ID AAU00507 standard; Protein; 380 AA.  
XX  
AC AAU00507;  
XX  
DT 29-AUG-2001 (first entry)  
XX  
DE E. coli heat-labile enterotoxin (LT) mutant Ldell110/112.  
XX  
KW Heat-labile enterotoxin; LT; Lts63Y; Ldell110/112; mutant;  
KW detoxified and immunologically active protein; ADP-ribosylation; Gs;  
KW endotoxin; diarrhoea; mutein.  
XX  
OS Escherichia coli strain K88ac.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT 7  
FT Misc\_feature /note= "Important residue for enzymatic activity"  
FT 44  
FT Misc\_feature /note= "Important residue for enzymatic activity"  
FT 58..72  
FT Region /note= "Forms the NAD-binding site"  
FT 61  
FT Misc\_feature /note= "Important residue for enzymatic activity"  
FT 257  
FT Misc-difference /note= "Encoded by TG"  
FT  
FT  
XX WO200119998-A1.  
XX  
XX  
PD 22-MAR-2001.  
XX  
PF 15-SEP-1999; 99WO-KR00555.  
XX  
PR 15-SEP-1999; 99WO-KR00555.  
XX  
XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.  
XX  
XX Park EJ, Kim JS, Chang J, Yum J, Chung S;  
XX  
XX WPI: 2001-281524/29.  
DR N-PSDB; AAS01506.  
XX  
XX New detoxified mutants of Escherichia coli heat-labile enterotoxin  
PT useful as vaccine for preventing and treating diarrhoea, and as adjuvant  
PT for antibody production  
XX  
XX Claim 6; Page 42-44; 48pp; English.  
XX  
XX The present sequence represents Escherichia coli heat-labile  
CC enterotoxin (LT) mutant Ldell110/112. Lts63Y (AAU00506) and Ldell110/112  
CC are two novel detoxified and immunologically active proteins (LT  
CC mutants) derived by site-directed mutagenesis of the A1 subunit of wild  
CC type LT. The substitution of Ser to Tyr at position 63 in Lts63Y blocks  
CC NAD-binding. Deletion of Glu residues at positions 110 and 112 in  
CC Ldell110/112 eliminate the enzymatic activity of LT. The A1 subunit of  
CC wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that  
CC regulates CAMP levels. The resulting increase in CAMP is the cause of  
CC diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities  
CC of mutant heat-labile endotoxins Lts63Y and Ldell110/112 were tested.  
CC Groups of mice were immunised with Lts623Y or Ldell110/112. The control  
CC faecal antibody titres to LT were determined. The results showed that  
CC mice immunised with Lts63Y or Ldell110/112 contained high and  
CC comparable level of anti-LT antibodies in sera and faecal extracts  
CC compared with those immunised with wild-type LT. The LT mutants are  
CC useful as a vaccine for preventing and treating diarrhoea and as an  
CC adjuvant for antibody production.  
XX  
XX Sequence 380 AA;

Query Match 98.2%; Score 1280; DB 22; Length 380;  
Best Local Similarity 98.8%; Pred. No. 1.6e-124;  
Matches 237; Conservative 1; Mismatches 0; Indels 2; Gaps 1;  
QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQNNINLYDHARTGTQTGFVRYDDGYV 60  
DB 19 ngdrlYradsrppdeikrsgglmpRGHNEYfdrGTqnnInlydhARTgtQTGFvRYddgyv 78  
QY 61 STSLSLRSALHLAGOSILSGYSTYYIYVIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120  
DB 79 stslsLrsahlagosilsgystyyIyVIatapnmfnvndvlgvysphpy--qvsalggip 136  
QY 121 YSQIYGYWYVNFVIDERLHNRREYRDYRYENLAPAEDEGYRLAGPPDHQAWRESPWI 180  
DB 137 ysqIygywYvNfViderlHnrreYrdYryeNlApaeDEgyRLagppDHqAwreSPwi 196  
QY 181 HHAPOGCGNSRRTITGTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIDEL 240  
DB 197 hhapogcgnsrRTITgtcNEetqnlStIylrEYqskvkrQIFsdYqseVDIynRIdeL 256  
RESULT 11  
AAV96650  
ID AAV96650 standard; Protein; 259 AA.  
XX  
AC AAV96650;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Plant-optimized E. coli LT-A A63K+R192G mutant.  
XX  
KW Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;  
KW adjuvant; anti-bacterial; A63K+R192G; double mutein.  
XX  
OS Escherichia coli.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..19  
FT /label= signal\_peptide  
FT Protein 20..259  
FT /label= mature\_protein  
FT Misc-difference 82  
FT /label= S63K  
FT /note= "the wild type serine is replaced by lysine  
FT at position 63 of the mature sequence"  
FT Misc-difference 211  
FT /label= R192G  
FT /note= "the wild type arginine is replaced by glycine  
FT at position 192 of the mature sequence"  
XX  
XX WO200037609-A2.  
XX  
XX 29-JUN-2000.  
XX  
XX 22-DEC-1999; 99WO-US30747.  
XX  
XX 22-DEC-1998; 98US-0113507.  
XX  
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.  
XX (NASO/) MASON H S.  
XX (ARNT/) ARNTZEN C J.  
XX  
XX Mason HS, Arntzen CJ;  
XX WPI: 2000-442653/38.  
XX  
XX New polynucleotides encoding LT-A or CT-A polypeptides for the  
PT transformation of plant cells, useful in immunogenic compositions to  
PT elicit immune responses in animals  
XX  
XX Example 5; Page -; 103pp; English.  
XX





ID AAW67772 standard; Protein; 237 AA.  
 AC AAW67772;  
 DT 11-MAY-1999 (first entry)  
 DE E. coli heat labile toxin.  
 XX  
 KW A subunit; heat labile toxin; ADP-ribosylation; mutant; detoxification;  
 KW parenteral adjuvant; antigen; antigen; immunisation; humoral response;  
 KW cell-mediated immune response; virus; bacterium; parasite; fungus;  
 KW tumour; allergen; pathogen; AIDS; autoimmune disease; cancer; antibody;  
 KW systemic lupus erythematosus; Alzheimer's disease; diagnosis.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO9842375-A1.  
 PD 01-OCT-1998.  
 XX  
 PF 19-MAR-1998; 98WO-0505454.  
 XX  
 PR 18-MAR-1998; 98US-0044696.  
 PR 21-MAR-1997; 97US-0041227.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 PI Barchfeld G, Del Giudice G, Rappuoli R;  
 XX  
 DR WPI; 1999-070064/06.  
 DR N-PSDB; AAV81595.  
 XX  
 PT Detoxified mutants of bacterial ADP-ribosylating toxins as  
 PT parenteral adjuvants - useful to enhance humoral and cell-mediated  
 PT immune responses in vertebrates when administered with selected  
 PT antigen e.g. in disease treatment  
 XX  
 PS Disclosure; Fig 1A-B; 51pp; English.  
 XX  
 CC This sequence corresponds to the amino acid sequence of the A subunit of  
 CC the E. coli heat labile toxin, an example of a bacterial ADP-ribosylating  
 CC toxin. A mutant detoxified form of this protein is used in a parenteral  
 CC adjuvant composition, which comprises the detoxified protein, at least  
 CC one selected antigen and optionally a pharmaceutically acceptable  
 CC (optionally topical) vehicle. The adjuvant composition can be  
 CC administered parenterally in conjunction with at least one antigen in  
 CC methods to immunise vertebrate subjects. The adjuvant has the ability  
 CC to enhance the humoral and cell-mediated immune responses elicited by  
 CC the antigen (e.g. by making the antigen more strongly immunogenic or  
 CC necessitating fewer/lower antigen doses). It can be administered  
 CC prior/subsequent to the antigen, and is preferably administered within  
 CC a short space of time to the same site; it can also be administered in  
 CC isolation from antigens as a boost following systemic or mucosal antigen  
 CC administration. Most preferably, the adjuvant is co-administered with  
 CC the antigen in the compositions and a pharmaceutically acceptable  
 CC carrier. The antigen may be derived from viruses, bacteria, parasites  
 CC and fungi or may be tumour antigens, self-antigens and allergens. The  
 CC compositions are therefore useful in the treatment and prevention of  
 CC e.g. viral diseases, allergic manifestations, diseases caused by  
 CC pathogens (e.g. bacteria or parasites), AIDS, autoimmune diseases  
 CC (e.g. Systemic Lupus Erythematosus), Alzheimer's disease and cancers.  
 CC The adjuvant can also be used to prepare antibodies against selected  
 CC antigen(s), useful e.g. for diagnostic purposes or for antigen  
 CC purification.  
 XX  
 SQ Sequence 237 AA;  
 XX  
 Query Match 91.3%; Score 1190.5; DB 20; Length 237;  
 Best Local Similarity 92.9%; Pred. No. 1.7e-115;  
 Matches 223; Conservative 3; Mismatches 11; Indels 3; Gaps 1;  
 QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 60

Db 1 ngdrllyradsrppdeikrslmprrgqneyfdrgtqnninlydhartqgtgfvyddgyv 60  
 QY 61 STSLSLRSAHLAGOSILSGYSTVIYVIATAPNNFNVDVLGVSPHPYEQEVSALGGIP 120  
 Db 61 stslslrsahlagqylsgysltiyvia---mmfnvndvisysphyeqevsalggip 117  
 QY 121 YSQIYGYRVNFGVIDRLHRNREYRDYRNINLNIAPAEQYRLAGFPDPHQAWREEPWI 180  
 Db 118 ysqilygyrvnfgvlderlhrnreayrdyrynlniapaedgyrlagfpdhdqawreepwi 177  
 QY 181 HHAPQCGGNSRRITGTCTCNEETQNLSTIYLREYQSKVKQIFSDYQSEVDIYNRIDEL 240  
 Db 178 hhapqcgdsrritgtdctneetqnlstiylyreyqskvrqifsdysqevdiynrirdel 237  
 RESULT 14  
 AAR38728  
 ID AAR38728 standard; Protein; 236 AA.  
 XX  
 AC AAR38728;  
 XX  
 DT 08-DEC-1993 (first entry)  
 XX  
 DE E.coli heat labile toxin subunit A.  
 XX  
 KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;  
 KW protomer A; site-directed mutagenesis; reduced toxicity;  
 KW ADP-ribosyltransferase activity.  
 XX  
 OS Escherichia coli.  
 XX  
 PN WO9313202-A.  
 XX  
 PD 08-JUL-1993.  
 XX  
 PF 30-DEC-1992; 92WO-EP03016.  
 XX  
 PR 31-DEC-1991; 91IT-OMI3513.  
 XX  
 PA (BIOC-) BIOGINE SCLAVO SPA.  
 XX  
 PI Domenighini M, Hol W, Pizza M, Rappuoli R;  
 XX  
 DR WPI; 1993-227320/28.  
 DR N-PSDB; AAQ42768.  
 XX  
 PT Immunogenic detoxified mutant cholera toxin and heat labile toxin  
 PT - useful as vaccines against infection by Vibrio cholerae and  
 PT enterotoxin producing Escherichia coli  
 XX  
 PS Disclosure; Fig 2; 60pp; English.  
 XX  
 CC This is the sequence of the A subunit of the heat labile toxin (LT-A)  
 CC of a strain of E.coli known to affect humans. The sequence was  
 CC published by Yamamoto et al, J.Biol. Chem., 259, 5037-5044.  
 CC Mutations at selected positions within this sequence have been found  
 CC to reduce toxicity (see AAR38730-R38732 and AAR44016-R44025). The  
 CC invention relates to such immunogenic, detoxified proteins and their  
 CC use in vaccines to protect against enterotoxigenic E.coli.  
 XX  
 SQ Sequence 236 AA;  
 XX  
 Query Match 90.7%; Score 1183; DB 14; Length 236;  
 Best Local Similarity 93.3%; Pred. No. 1e-114;  
 Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;  
 QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 60  
 Db 1 ngdrllyradsrppdeikrslmprrgqneyfdrgtqnninlydhartqgtgfvyddgyv 59  
 QY 61 STSLSLRSAHLAGOSILSGYSTVIYVIATAPNNFNVDVLGVSPHPYEQEVSALGGIP 120

Db 60 stslsrsahlagqylsgysltiyvia---nmfnvndvisvysphpyeqeysalggip 116  
 QY 121 YSQIYGYRVNFGVIDERLHNRNREYRDYRNINIAPAEDGYRLAGFPDPDQAWREPEWI 180  
 Db 117 ysqlygyrvnfgvlderlhrnreayrdyrynlniapaedgyrlagfpdqawreepwi 176  
 QY 181 HHAPQCGNSRRTITGTCNEETQNLSTIYLREYQSKVKRQIFSDYOSEVDIYNRIDEL 240  
 Db 177 hhapqcgdsrstitgtctneetqnlstlylreygskvkqrifsdysqevdiynrirdel 236

## RESULT 15

AAU14105  
 ID AAU14105 standard; peptide; 254 AA.

XX AC AAU14105;

XX 21-NOV-2001 (first entry)

XX Peptide sequence from Escherichia coli heat labile enterotoxin A.

XX Anti-retroviral; DP178-like; DP107-like; heat labile enterotoxin A;  
 KW antifusogenic; antiviral; HIV transmission.

XX OS Escherichia coli.

XX PN WO200151673-A2.

XX PD 19-JUL-2001.

XX PF 05-JUL-2000; 2000WO-US35727.

XX PR 09-JUL-1999; 99US-0350841.

XX PA (TRIM-) TRIMERIS INC.

XX PI Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;

XX DR WPI; 2001-442157/47.

XX Identifying a compound that inhibits the formation of or disrupts a  
 PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral  
 PT or intracellular modulatory activity, by detecting the formation of a  
 XX DP107/DP178 complex -

PS Disclosure; Fig 43; 259pp; English.

XX The present invention relates to peptides which exhibit anti-retroviral  
 CC activity. The peptides of the invention (AAU12559-AAU14009) comprise  
 CC DP178-like and DP107-like peptides. The DP178 peptide corresponds  
 CC to amino acids 639-673 of the transmembrane protein gp41 from human  
 CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide  
 CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention  
 CC also relates to a method of identifying compounds that inhibit the  
 CC formation of or disrupts a DP107/DP178 complex. The method comprises  
 CC detecting the formation of a DP107/DP178 complex, both in the presence  
 CC or absence of a test compound, in a reaction mixture containing DP107  
 CC and DP178 peptides. The method is useful for identifying compounds,  
 CC including small molecule compounds, which may themselves exhibit  
 CC antifusogenic, antiviral or intracellular modulatory activity. The  
 CC DP178-like/DP107-like peptides are useful to inhibit human and non-human  
 CC retroviral, particularly HIV, transmission to uninfected cells. The  
 CC present sequence represents a peptide sequence from Escherichia coli  
 CC heat labile enterotoxin A.

XX Sequence 254 AA;

Query Match

Best Local Similarity 93.3%; Score 1183; DB 22; Length 254;

Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

QY \* 1 NGDMLYRADSRPPDEIKRSGGLMPRGHNEYFDRCGTQMNINLYDHARTGTGTFVRYDDGYV 60  
 Db 19 ngdrlyradsrppdpeikrfrslmprg-neyfdrgtqmninlydhargtqtfvryddgyv 77  
 QY 61 STSLSLSAHLAQSILSGYSTIYIYIATAPNMFNVDVLGYVSPHPYEQEVSALGGIP 120  
 Db 78 stslsrsahlagqylsgysltiyvia---nmfnvndvisvysphpyeqeysalggip 134  
 QY 121 YSQIYGYRVNFGVIDERLHNRNREYRDYRNINIAPAEDGYRLAGFPDPDQAWREPEWI 180  
 Db 135 ysqlygyrvnfgvlderlhrnreayrdyrynlniapaedgyrlagfpdqawreepwi 194  
 QY 181 HHAPQCGNSRRTITGTCNEETQNLSTIYLREYQSKVKRQIFSDYOSEVDIYNRIDEL 240  
 Db 195 hhapqcgdsrstitgtctneetqnlstlylreygskvkqrifsdysqevdiynrirdel 254

Search completed: June 10, 2002, 17:31:46

Job time: 1053 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:29:43 ; Search time 19.04 Seconds  
(without alignments)  
1211.210 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 1304  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNRIQDEL 240

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1287	98.7	258	1 QLECA	heat-labile entero
2	1088	83.4	258	1 XVVCA	cholera enterotoxi
3	679.5	52.1	259	2 A29831	heat-labile entero
4	394.5	30.3	125	2 A61345	heat-labile entero
5	152	11.7	46	2 S39239	cholera enterotoxi
6	96	7.4	269	1 B25973	pertussis toxin ch
7	93	7.1	268	1 WEBR11	pertussis toxin ch
8	93	7.1	269	1 WEBR1P	pertussis toxin ch
9	91.5	7.0	472	1 A47402	fatty acid binding
10	91	7.0	392	2 T01849	hypothetical prote
11	90.5	6.9	702	2 A61619	arylphorin precurs
12	88.5	6.8	332	2 S60935	hypothetical prote
13	88	6.7	269	1 A25973	pertussis toxin ch
14	88	6.7	547	2 T43666	hypothetical prote
15	87	6.7	809	2 T40574	guanine nucleotide
16	86.5	6.6	145	2 G81094	hypothetical prote
17	86.5	6.6	566	2 A48860	beta-glucosidase,
18	86	6.6	423	2 A72616	probable glutamate
19	85.5	6.6	435	2 E72563	probable tRNA nucl
20	83.5	6.4	870	2 S27514	mosquitocidal toxi
21	83	6.4	254	2 A34759	prion protein - Ch
22	83	6.4	254	2 B34759	prion protein - go
23	83	6.4	353	2 F96903	hypothetical prote
24	83	6.4	7829	2 T15789	hypothetical prote
25	82	6.3	256	2 JU0268	major prion protei
26	82	6.3	446	1 S74920	fmu protein - Syne
27	82	6.3	469	2 S77704	6-phosphofructo-2-
28	82	6.3	471	1 KIITFB	6-phosphofructo-2-
29	81.5	6.2	828	1 D39142	outer membrane ush

30	81	6.2	264	2 A54330	major prion protei
31	81	6.2	899	2 B38529	nikB protein - Esc
32	80.5	6.2	173	2 T41773	LEF-6 orf28 - Bomb
33	80.5	6.2	252	2 E71704	hypothetical prote
34	80.5	6.2	880	2 S60137	beta-N-acetylhexos
35	80	6.1	254	1 UJHVIH	major prion PrP-Sc
36	80	6.1	257	2 A23545	major prion PrP27-
37	80	6.1	260	2 S53629	major prion protei
38	80	6.1	264	2 S37137	prion protein - gr
39	80	6.1	3972	2 S75251	hypothetical prote
40	79	6.1	474	2 S16250	phytoene dehydroge
41	79	6.1	509	2 AD0649	probable secreted
42	79	6.1	630	2 A86389	70-3K hypothetical
43	78.5	6.0	648	2 S10869	enterotoxin A - Cl
44	78.5	6.0	767	2 T43398	SCF complex protei
45	78.5	6.0	983	2 H72510	probable ribonucle

ALIGNMENTS

RESULT 1

QLECA

heat-labile enterotoxin A precursor - Escherichia coli

C:Species: Escherichia coli

C>Date: 30-Apr-1981 #sequence\_revision 17-Oct-1997 #text\_change 18-Jun-1999

C:Accession: I55231; A01817; A26946

R:Yamamoto, T.; Tamura, T.; Yokota, T.

J. Biol. Chem. 259, 5037-5044, 1984

A:Title: Primary structure of heat-labile enterotoxin produced by Escherichia coli pa

A:Reference number: I55231; MUID:84185610

A:Accession: I55231

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-258 <RES>

A:Cross-references: GB:K01995; NID:g148027; PIDN:AAA24685.1; PID:g148028

R:Spicer, E.K.; Noble, J.A.

J. Biol. Chem. 257, 5716-5721, 1982

A:Title: Escherichia coli heat-labile enterotoxin. Nucleotide sequence of the A subun

A:Reference number: A01817; MUID:82167425

A:Accession: A01817

A:Molecule type: DNA

A:Residues: 1-21, 'R', '23-36, 'FRS', '40-44, 46-92, 'Y', '94-99, 'LTIYI', '105-107, 111-118, 'IS', '1

A:Cross-references: EMBL:V00275; NID:g41339; PIDN:CAA23532.1; PID:g41340

A:Note: the authors translated the codon TAT for residue 93 as Ser

R:Yamamoto, T.; Gojobori, T.; Yokota, T.

J. Bacteriol. 169, 1352-1357, 1987

A:Title: Evolutionary origin of pathogenic determinants in enterotoxigenic Escherichi

A:Reference number: A26946; MUID:87137303

A:Accession: A26946

A:Molecule type: DNA

A:Residues: 1-21, 'R', '23-206, 'N', '208-230, 'E', '232-255, 'D', '257-258 <YAM>

A:Cross-references: EMBL:M15363

C:Comment: The heat-labile enterotoxin molecule contains one A chain and five or six

C:Genetics:

A:Gene: elta

C:Superfamily: heat-labile enterotoxin chain A

C:Keywords: enterotoxin

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-258/Product: heat-labile enterotoxin chain A #status predicted <MAT>

Query Match 98.7%; Score 1287; DB 1; Length 258;

Best Local Similarity 98.3%; Pred. No. 1.3e-110;

Matches 236; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEVFDRTGTMNINLYDHARGTQGFVYDDGYV 60

Db 19 NGDKLLRADSRPPDEIKRSGGLMPRGHNEVFDRTGTMNINLYDHARGTQGFVYDDGYV 78

QY 61 STSLSLRSARLAGQSILSGYSTYIYVIATAPNMFNVDVLGYVSPRPYEQEYSALGGIP 120

Db 79 STSLSLSAHLAQSGILSGYSTYYIVATAPNMFNVNDVLGVSPHPYEQEVSALGGIP 138

QY 121 YSQIYGYRVNFGVIDERLHRNREYRDYRNLIAPAEDGYRLAGFPDPHQAWREEPWI 180  
|||||

Db 139 YSQIYGYRVNFGVIDERLHRNREYRDYRNLIAPAEDGYRLAGFPDPHQAWREEPWI 198  
|||||

QY 181 HHAPOGCGNSRITGDTCTNEETONLSTIYLRYSQSKVRQIFSDYQSEVDIYNRIDEL 240  
|||||

Db 199 HHAPOGCGDSRITGDTCTNEETONLSTIYLRYSQSKVRQIFSDYQSEVDIYNRIDEL 258  
|||||

RESULT 2

XVUCA  
cholera enterotoxin chain A precursor VC1457 [validated] - Vibrio cholerae  
C:Species: Vibrio cholerae  
C:Date: 06-Jul-1982 #sequence\_revision 26-Jan-1996 #text\_change 01-Sep-2000  
A:Accession: A05129; S14623; S14625; A91286; A91746; A92298; S17665; B43864; A82  
R:Nekalanos, J.J.; Swartz, D.J.; Pearson, G.D.N.; Harford, N.; Groyne, F.; de Wilde, M.  
Nature 308, 551-557, 1983  
A:Reference number: A93320; MUID:84068199  
A:Accession: A05129  
A:Molecule type: DNA  
A:Residues: 1-258 <MEK>  
R:Dams, E.; de Wolf, M.; Dierick, W.  
submitted to the EMBL Data Library, March 1991  
A:Description: Correction of the cholera toxin nucleotide sequence of the Vibrio cholera  
A:Reference number: S14623  
A:Accession: S14623  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <DAL>  
A:Cross-references: EMBL:X58786; NID:g48420; PIDN:CAA1592.1; PID:g48421  
A:Experimental source: strain 2125  
A:Accession: S14625  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <DA2>  
A:Cross-references: EMBL:X58785; NID:g48888; PIDN:CAA1590.1; PID:g48889  
A:Experimental source: strain 569B  
R:Lai, C.Y.; Cancedda, F.; Chang, D.  
FEBS Lett. 100, 85-89, 1979  
A:Title: Primary structure of cholera toxin subunit A-1. Isolation, partial sequences an  
A:Reference number: A91268; MUID:79169830  
A:Accession: A91268  
A:Molecule type: protein  
A:Residues: 37-38,'L',40-44,'SE',47-49,'B',51-55,'B',57,'B',59-60,'B',62-66,'Z',68-72,'I  
R:Duffy, L.K.; Peterson, J.W.; Kurosky, A.  
FEBS Lett. 126, 187-190, 1981  
A:Title: Isolation and characterization of a precursor form of the 'A' subunit of cholera  
A:Reference number: A91286; MUID:81212799  
A:Accession: A91286  
A:Molecule type: protein  
A:Residues: 19,'N',21-27 <DUF>  
R:Klapper, D.G.; Finkelstein, R.A.; Capra, J.D.  
Immunochimistry 13, 605-611, 1976  
A:Title: Subunit structure and N-terminal amino acid sequence of the three chains of cho  
A:Reference number: A91746; MUID:76259136  
A:Accession: A91746  
A:Molecule type: protein  
A:Residues: 19-36,'R',38;213-232 <KLA>  
R:Duffy, L.K.; Peterson, J.W.; Kurosky, A.  
J. Biol. Chem. 256, 12252-12256, 1981  
A:Title: Covalent structure of the gamma chain of the A subunit of cholera toxin.  
A:Reference number: A92298; MUID:82053094  
A:Accession: A92298  
A:Molecule type: protein  
A:Residues: 213-246,'ID',249-255,'N',257-258 <DU2>  
R:Dams, E.; de Wolf, M.; Dierick, W.  
Biochim. Biophys. Acta 1090, 139-141, 1991  
A:Title: Nucleotide sequence analysis of the CT operon of the Vibrio cholerae classical  
A:Reference number: S17665; MUID:91355224  
A:Accession: S17665  
A:Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-258 <DAM>  
A:Cross-references: EMBL:X58785; NID:g48888; PIDN:CAA1590.1; PID:g48889  
R:Baudry, B.; Fasanio, A.; Ketley, J.; Kaper, J.B.  
Infect. Immun. 60, 428-434, 1992  
A:Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.  
A:Reference number: A43864; MUID:92112300  
A:Accession: B43864  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-6 <BAU>  
A:Cross-references: GB:M83563; NID:gl55314; PIDN:AAA27583.1; PID:gl55316  
A:Note: sequence extracted from NCBI backbone (NCBIN:77488, NCBI:77496)  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers  
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: A82197  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <HEI>  
A:Cross-references: GB:AE004224; GB:AE003852; NID:g9655952; PIDN:AAF94614.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1457  
A:Map position: 1  
C:Complex: the cholera enterotoxin molecule contains three kinds of chains; an alpha  
ciate noncovalently with the subunit B, an aggregate of five beta chains  
C:Function:  
A:Description: the active component of the toxin that is primarily responsible for s  
.2.5) activity also activates intracellular adenylyl cyclase  
C:Superfamily: heat-labile enterotoxin chain A  
C:Keywords: enterotoxin  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-212/Product: cholera enterotoxin alpha chain #status experimental <CTA>  
F:213-258/Product: cholera enterotoxin gamma chain #status experimental <CTG>  
F:217/Disulfide bonds: Interchain (to alpha chain) #status predicted

Query Match 83.4%; Score 1088; DB 1; Length 258;  
Best Local Similarity 81.7%; Pred. NO. 2.4e-92;  
Matches 196; Conservative 24; Mismatches 20; Indels 0; Gaps 0;

QY 1 NGDRLYRADSRPDEIKRSGGLMPRGHNEYFDRTGTMNINLYDHARGTQTGFVRDDGVV 60  
|||||

Db 19 NDDKLYRADSRPDEIKSGGLMPRGQSEYFDRTGTMNINLYDHARGTQTGFVRDDGVV 78  
|||||

QY 61 STSLSLSAHLAQSGILSGYSTYYIVATAPNMFNVNDVLGVSPHPYEQEVSALGGIP 120  
|||||

Db 79 STSLSLSAHLVQGITLSGHSTYYIVATAPNMFNVNDVLGAYSPHPDEQVSALGGIP 138  
|||||

QY 121 YSQIYGYRVNFGVIDERLHRNREYRDYRNLIAPAEDGYRLAGFPDPHQAWREEPWI 180  
|||||

Db 139 YSQIYGYRVNFGVIDERLHRNREYRDYRNLIAPAADGYRLAGFPDPHQAWREEPWI 198  
|||||

QY 181 HHAPOGCGNSRITGDTCTNEETONLSTIYLRYSQSKVRQIFSDYQSEVDIYNRIDEL 240  
|||||

Db 199 HHAPOGCGNAPRSMNSNTCDKTSGLGVKFLDEYQSKVRQIFSGYQSDIDTHNRKIDEL 258  
|||||

RESULT 3

A29831  
heat-labile enterotoxin IIA chain A precursor - Escherichia coli  
N:Alternate names: LT-IIa  
C:Species: Escherichia coli  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 16-Jul-1999  
A:Accession: A29831  
R:Pickett, C.L.; Weinstein, D.L.; Holmes, R.K.  
J. Bacteriol. 169, 5180-5187, 1987  
A:Title: Genetics of type IIA heat-labile enterotoxin of Escherichia coli: operon fus  
A:Reference number: A91849; MUID:88032841

A;Accession: A29831  
A;Molecule type: DNA  
A;Residues: 1-259 <PIC>  
A;Cross-references: GB:M17894; NID:g146671; PIDN:AAA24093.1; PID:g146672  
A;Note: the authors translated the codon TAT for residue 225 as Thr  
C;Superfamily: heat-labile enterotoxin chain A  
F;1-18/Domain: signal sequence #status predicted <SIG>  
F;19-259/Product: heat-labile enterotoxin Iia chain A #status predicted <LTA>

Query Match 52.1%; Score 679.5; DB 2; Length 259;  
Best Local Similarity 54.6%; Pred. No. 8.2e-55;  
Matches 130; Conservative 39; Mismatches 66; Indels 3; Gaps 2;

QY 6 YRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTGTGFVRDGDGYVSTLS 65  
DB 22 FRADSRPPDEIRAGGLLPGRQOAYERGTPINLYEHARGTGTGTRVNDGYVSTVT 81

QY 66 LRSALAGQSILSGSYTYIYVIATAPNMFVNDVGLVSPHPYQEVSAALGGIPYSQIY 125  
DB 82 LRQHLIGQNLGILSYNEYIYVAPAPNLFVNGVGLGRYSPYSPSENEFAALGGIPLSQII 141

QY 126 GWYRVNGVIDERLHRNREYRDYRNINLAPADGYRLAGFPDPHQAWEERPHIHAPO 185  
DB 142 GWYRVSEGAIEGGQMRDYGDLFRGLTVAPNEDGYQLAGFSPNFPAREMPWSTEPE 201

QY 186 GCGNSRTITGDCNEETQNLSTIYLREYQSKVKROI-FSDYQSEVDI--YNRIRDEL 240  
DB 202 QCVPNNREFGGVCISATNVLKSLVDLMFKLLKRRLLATFFSEDDFIGVHGERDEL 259

RESULT 4  
A61345  
heat-labile enterotoxin A chain precursor - Escherichia coli (fragments)  
C;Species: Escherichia coli  
C;Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 17-Mar-1999  
R;Spicer, E.K.; Kavanaugh, W.M.; Dallas, W.S.; Falkow, S.; Konigsberg, W.H.; Schafer, D.  
C;Accession: A61345  
Proc. Natl. Acad. Sci. U.S.A. 78, 50-54, 1981  
A;Title: Sequence homologies between a subunits of Escherichia coli and Vibrio cholerae  
A;Reference number: A61345; MUID:81223767  
A;Accession: A61345  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-125 <SPI>  
A;Cross-references: GB:K00433  
A;Note: authors translated the codon GAG for residue 27 as Asp, CAA for residue 53 as Gl  
C;Superfamily: heat-labile enterotoxin chain A

Query Match 30.3%; Score 394.5; DB 2; Length 125;  
Best Local Similarity 39.9%; Pred. No. 4.7e-29;  
Matches 87; Conservative 2; Mismatches 6; Indels 123; Gaps 2;

QY 1 NGBRLRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTGTGFVRDGDYV 60  
DB 19 NGBRLYRAESRPPDEIKRFRSLPRG-NEYFDRGTOMNINLYDHARGTGTGFVR----- 71

QY 61 STSLSLSAHLAQSIILSGSYTYIYVIATAPNMFVNDVGLVSPHPYQEVSAALGGIP 120  
DB 72 ----- 71

QY 121 YSQIYGWYRVNGVIDERLHRNREYRDYRNINLAPADGYRLAGFPDPHQAWEERPHI 180  
DB 72 -----RTWI 75

QY 181 HHAPQCGNSRRTITGDCNEETQNLSTIYLREYQSKV 218  
DB 76 HHAPQCGDSRRTITGDCNEETQNLSTIYLREYQSKV 113

RESULT 5  
S39239

cholera enterotoxin chain A2 - Vibrio cholerae (serotype 0139)  
N;Alternate names: CTA2 cholera toxin chain A2  
C;Species: Vibrio cholerae  
A;Variety: serotype 0139  
C;Date: 08-May-1995 #sequence\_revision 26-Jul-1996 #text\_change 16-Jul-1999  
C;Accession: S39239; S39240  
R;Lebens, M.; Holmgren, J.  
submitted to the EMBL Data Library, November 1993  
A;Description: Structure and arrangement of the Cholera toxin genes in vibrio Cholera  
A;Reference number: S39238  
A;Accession: S39239  
A;Molecule type: DNA  
A;Residues: 1-46 <LEP>  
A;Cross-references: EMBL:X76390; NID:g433856; PIDN:CAA53974.1; PID:g433858; EMBL:X763  
A;Experimental source: strain 4260B; serotype 0139  
C;Superfamily: heat-labile enterotoxin chain A

Query Match 11.7%; Score 152; DB 2; Length 46;  
Best Local Similarity 63.6%; Pred. No. 2.4e-07;  
Matches 28; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 197 DTCNEETQNLSTIYLREYQSKVKROI-FSDYQSEVDIYNRIRDEL 240  
DB 3 NTCDEKTSGLGVKFLDEYQSKVKRQYFSGYQSDIDTHNRIRDEL 46

RESULT 6  
B25973  
pertussis toxin chain S1 precursor - Bordetella parapertussis  
C;Species: Bordetella parapertussis  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: B25973  
R;Arico, B.; Rappuoli, R.  
J. Bacteriol. 169, 2847-2853, 1987  
A;Title: Bordetella parapertussis and Bordetella bronchiseptica contain transcription  
A;Reference number: A25973; MUID:6722217  
A;Accession: B25973  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-269 <ARI>  
C;Superfamily: pertussis toxin chain S1

Query Match 7.4%; Score 96; DB 1; Length 269;  
Best Local Similarity 24.2%; Pred. No. 0.34; Mismatches 24; Indels 90; Gaps 12;

QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTGTGFVRDGDGYVSTSL 64  
DB 41 VYRYDSRPPEDVFQNGFTAGNND-----NVLEHLTGRSCQVSGSSNAFVSTSS 89

QY 65 SLR-----SAHLAQSIILSGSYTYIYVIATAPNMFVNDVGLVSPHPYE 110  
DB 90 SRRYTEVLEHRMOEAVEAERAGRG--TGHFIGIYEI-RADNRF----- 131

QY 111 QEVSAIGIPYSQIYGWYRVNGVIDERLHRNREYRDYRNINL-IAPAEDGYRLAGFP 168  
DB 132 -----YGAASSYF-----EYVDYGDNAGRILAGA-----LATYQ 161

QY 169 PDHQAWEERPHIHAPOCGNSRRT-----ITGDCNEETQNLSTIYLREYQSKVKROI 224  
DB 162 SEYLAHRRIP-----PENIRTVRVYHNGITGETTTEYPNL-----RYVSQQTTRAN 210

QY 225 DYQS 228  
DB 211 PYTS 214

RESULT 7  
WEBRI1  
pertussis toxin chain S1 precursor (version 1) - Bordetella pertussis  
N;Alternate names: histamine-sensitizing factor; islet-activating protein; leukocytos

N:Contains: NAD+ ADP-ribosyltransferase (EC 2.4.2.30)

C:Species: Bordetella pertussis

C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 24-Sep-1999

C:Accession: A24394

R:Locht, C.; Keith, J.M.

Science 232, 1258-1264, 1986

A:Title: Pertussis toxin gene: nucleotide sequence and genetic organization.

A:Reference number: A94285; MUID:86208173

A:Accession: A24394

A:Molecule type: DNA

A:Residues: 1-268 <LOC>

A:CROSS-references: GB:M13223

C:Comment: Pertussis toxin contains five different chains, S1-S5. These are organized in to the membrane receptors. Dimers of S2-S4 and S3-S4 are held together by S5. The genes to Comment: Pertussis toxin chains are irreversible uncoupling of the regulatory GTP-binding protein. The substrates for the ADP-ribosyltransferase activity are the GTP-binding protein. Comment: Pertussis toxin is the major virulence factor of B. pertussis.

C:Comment: See PIR:WEBRIP for version 2 of this sequence.

C:Superfamily: pertussis toxin chain S1

C:Keywords: glycosyltransferase; mitogen; pentosyltransferase; toxin; virulence factor;

F;1-34/Domain: signal sequence #status predicted <SIG>

F;35-268/Product: pertussis toxin chain S1 #status predicted <MAT>

Query Match 7.1%; Score 93; DB 1; Length 268;

Best Local Similarity 24.0%; Pred. No. 0.63;

Matches 53; Conservative 22; Mismatches 62; Indels 84; Gaps 11;

QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRTQMNNINLYDHARGTQTGFVRYDDGYVSTSL 64

Db 41 VYRYSRPPDEVFQN-GFTAGNND-----NVLDLTLGRSCQVGSNSAFVSTSS 89

QY 65 SLR-----SAHLAQSILSGYSTYIYIATAPNMFNVNDVLGVSPHYE 110

Db 90 SRRYTEVYLEHRMOWEAEVARGG--TGHFIGYIEV-RADNMF-----LATYQ 131

QY 111 QEVSALGGIPYSQIYGYRYNFGVIDERLHRNREYRDYRYNRLN--IAPAEQYRLAGFP 168

Db 132 -----YGAASSYF-----EYVDYDGNAGRILAGA-----LATYQ 161

QY 169 PDQAWREEPWIHAPOGCGNSRT-----ITGDTCEETQN 205

Db 162 SEYLAHRIP-----PENIRRTVRYNHGTTTTEYSN 197

RESULT 8

WEBRIP

pertussis toxin chain S1 precursor (version 2) - Bordetella pertussis

N:Alternate names: histamine-sensitizing factor; islet-activating protein; lymphocytosis

C:Species: Bordetella pertussis

C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 24-Sep-1999

C:Accession: A24144; S06588; A36176; S04494

R:Nicoia A.; Perugini, M.; Franzini, C.; Casaglini, M.C.; Borri, M.G.; Antoni, G.; Almon

Proc. Natl. Acad. Sci. U.S.A. 83, 4631-4635, 1986

A:Title: Cloning and sequencing of the pertussis toxin genes: operon structure and gene

A:Reference number: A94104; MUID:86259651

A:Accession: A24144

A:Molecule type: DNA

A:Residues: 1-269 <NIC>

A:CROSS-references: GB:M14378; NID:gl44070; PIDN:AAA93980.1; PID:gl44071

R:Loosmore, S.M.; Cunningham, J.D.; Bradley, W.R.; Yao, F.L.; Dekaban, G.A.; Klein, M.H.

Nucleic Acids Res. 17, 8365, 1989

A:Title: A unique sequence of the Bordetella pertussis toxin operon.

A:Reference number: S06588; MUID:90045957

A:Accession: S06588

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-67, 'E', 69-231, 'Y', 233-269 <LOO>

A:CROSS-references: EMBL:X16347; NID:g39769; PIDN:CAA34397.1; PID:g39770

R:Cieplak, W.; Burnette, W.N.; Mar, V.L.; Kaljot, K.T.; Morris, C.F.; Chen, K.K.; Sato,

Proc. Natl. Acad. Sci. U.S.A. 85, 4667-4671, 1988

A:Title: Identification of a region in the S1 subunit of pertussis toxin that is require

A:Reference number: A36176; MUID:88263009

A:Accession: A36176

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: DNA

A:Residues: 35-64 <CIE>

R:Cockle, S.A.

FEBS Lett. 249, 329-332, 1989

A:Title: Identification of an active-site residue in subunit S1 of pertussis toxin by

A:Reference number: S04494; MUID:89290006

A:Accession: S04494

A:Molecule type: protein

A:Residues: 152-162, 'X', 164-166 <COC>

A:Experimental source: strain 10536

C:Comment: See PIR:WEBRIP for version 1 of this sequence.

C:Superfamily: pertussis toxin chain S1

C:Keywords: glycosyltransferase; mitogen; pentosyltransferase; toxin; virulence facto

F;1-34/Domain: signal sequence #status predicted <SIG>

F;35-269/Product: pertussis toxin chain S1 #status predicted <MAT>

Query Match 7.1%; Score 93; DB 1; Length 269;

Best Local Similarity 24.0%; Pred. No. 0.63;

Matches 53; Conservative 22; Mismatches 62; Indels 84; Gaps 11;

QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRTQMNNINLYDHARGTQTGFVRYDDGYVSTSL 64

Db 41 VYRYSRPPDEVFQN-GFTAGNND-----NVLDLTLGRSCQVGSNSAFVSTSS 89

QY 65 SLR-----SAHLAQSILSGYSTYIYIATAPNMFNVNDVLGVSPHYE 110

Db 90 SRRYTEVYLEHRMOWEAEVARGG--TGHFIGYIEV-RADNMF-----LATYQ 131

QY 111 QEVSALGGIPYSQIYGYRYNFGVIDERLHRNREYRDYRYNRLN--IAPAEQYRLAGFP 168

Db 132 -----YGAASSYF-----EYVDYDGNAGRILAGA-----LATYQ 161

QY 169 PDQAWREEPWIHAPOGCGNSRT-----ITGDTCEETQN 205

Db 162 SEYLAHRIP-----PENIRRTVRYNHGTTTTEYSN 197

RESULT 9

A47402

fatty acid binding/transport protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 19-May-1995 #sequence\_revision 12-Apr-1996 #text\_change 22-Jun-1999

C:Accession: A47402

R:Abumrad, N.A.; El-Maghrabi, M.R.; Amri, E.Z.; Lopez, E.; Grimaldi, P.A.

J. Biol. Chem. 268, 17665-17668, 1993

A:Title: Cloning of a rat adipocyte membrane protein implicated in binding or transpo

A:Reference number: A47402; MUID:93352566

A:Accession: A47402

A:Molecule type: mRNA

A:Residues: 1-472 <ABU>

A:CROSS-references: GB:I19658; NID:g310112; PIDN:AAA02878.1; PID:g310113

C:Superfamily: lysosomal integral membrane protein II

C:Keywords: glycoprotein; transmembrane protein

F;2-6/Domain: intracellular #status predicted <CYT1>

F;7-30/Domain: transmembrane #status predicted <TML>

F;31-439/Domain: extracellular #status predicted <EXT>

F;440-466/Domain: transmembrane #status predicted <TM2>

F;467-472/Domain: intracellular #status predicted <CYT2>

F;79,102,134,205,220,235,247,417/Binding site: carbohydrate (Asn) (covalent) #status

Query Match 7.0%; Score 91.5; DB 1; Length 472;

Best Local Similarity 20.1%; Pred. No. 1.8;

Matches 56; Conservative 39; Mismatches 85; Indels 99; Gaps 12;

QY 2 GDRLYRA----DSRPPDEIKRSGG---LMPRGHNEYFDRTQMNNINLYDHARGTQTGFVR 54

Db 58 GTTVYQFVDFQNFEEVAKNSKIKVIQGYTYRVYLAKE-NITQDPKDSVTSFVQ 116



QY 55 YDDGYVSTSLRSALHAGOSILSGYSTYYIYVIATAPNMFVNDVLGVYS----- 105  
Db 117 PNGAIFPSLSVGT-----NDNTVLNLAAVAAPHIYTNFSVOGLNLSLIKSKSS 168  
QY 106 -----PHPYEQEVSALGGIPYSQ-IYGYWRVNFV-----V 134  
Db 169 MFQTRSLKELLNGYKDPFLSLVPPYSTVGVF--YPYNTVDGVYKVSNGKDNISKVAI 226  
QY 135 IDE-RLHRRNEYDRYRNLNTAPADGYRLAGFPDPDQHWREEPWIHAPQCGGSSRT 193  
Db 227 IDTYKGRNLSYWEYCDMIN-----GTDAASFPP-----LGEKSRT 263  
QY 194 ITGDTCTNEETQNLSTIYLREYQSKVKROIIFSQYQSEVDI 232  
Db 264 -----LRFSSDICRSIYAVFESEVNL 285  
RESULT 10  
T01849  
hypothetical protein F9D12.5 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 26-Feb-1999 #sequence\_revision 26-Feb-1999 #text\_change 26-Feb-1999  
C:Accession: T01849  
R:Murray, J.; Langston, Y.; Ahrens, C.  
submitted to the EMBL Data Library, July 1998  
A:Description: The sequence of Arabidopsis thaliana F9D12.  
A:Reference number: Z1444  
A:Accession: T01849  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-392 <MUR>  
A:Cross-references: EMBL:AF077407; NID:g3319339; PID:g3319347  
C:Genetics:  
A:Map position: 4  
A:Introns: 36/3; 134/2; 191/1; 283/2; 339/1  
A:Note: F9D12.5

Query Match 7.0%; Score 91; DB 2; Length 392;  
Best Local Similarity 21.5%; Pred. No. 1.6;  
Matches 45; Conservative 44; Mismatches 76; Indels 44; Gaps 12;  
QY 29 EYDRTGTMNIN--LYDHARGTGTGFRYDDGVYSTSLRSALHAGQST--LSGYSTY- 83  
Db 162 EYLPGRWEVNVDLKFIH-NGKLNKYLAIISDGTGLKLYNDAKRWGFGQLIPHVFITYG 220  
QY 84 YIYVIATAPNMFVNDVLGVYSPHPYEQEVSALGGIPYSQIYGYWRVNFVGVIDERLHRRN 143  
Db 221 YI-----EQDIGSFGAEIFIVKPAQOQEKVTFISN-PPTNVFTWKILHFSILEDKFIYSD 274  
QY 144 EY--RDYRYRNLNIAEDG-----YRLA--GFPDPDQHWREEPWIHAPQCGGSSRTIT 195  
Db 275 DFLVEDRYWR-LGFNPKDGGGRPYALPTFLFAQGHKA-----NAVATNT 318  
QY 196 GDTCTNEETQNLSTIYLREYQSKVKROIIFS 224  
Db 319 WGAAN-----LRKNQRSTNHRQIYT 339

RESULT 11  
A61619  
arylphorin precursor - greater wax moth  
N:Alternate names: phenylalanine-rich protein Lhp76  
C:Species: Galleria mellonella (greater wax moth)  
C>Date: 21-Jul-1995 #sequence\_revision 11-Aug-1995 #text\_change 24-Sep-1999  
C:Accession: A61619  
R:Mummel, N.A.; Trevitt, P.M.; Silhacek, D.L.; Kumaran, A.K.  
Insect Biochem. Mol. Biol. 22, 333-342, 1992  
A:Title: Nucleotide sequence and structure of the arylphorin gene from Galleria mellonella  
A:Reference number: A61619  
A:Accession: A61619  
A:Molecule type: DNA  
A:Residues: 1-702 <MEM>

A:Cross-references: GB:M73793; NID:g159077; PIDN:AAA74229.1; PID:g159078  
A:Note: In the authors' translation, residue 306-Leu is omitted, residue 317-Asn is d  
C:Genetics:  
A:Introns: 30/1; 74/3; 352/3; 410/2  
A:Note: single copy gene  
C:Superfamily: arylphorin  
C:Keywords: glycoprotein; hemolymph; storage protein  
F:16/Domain: signal sequence #status predicted <SIG>  
F:211.481/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 6.9%; Score 90.5; DB 2; Length 702;  
Best Local Similarity 20.4%; Pred. No. 3.8;  
Matches 55; Conservative 37; Mismatches 82; Indels 95; Gaps 12;  
QY 5 LYRADSRPPDEIKRSGGLMPRGHN---EYEDRG-----TQMNINLYDHARGTGTG 53  
Db 149 IHRADTR-----GIVLPAPYEINPEYFMNSDLSKIYRIQMKGLIIPQGGPYGIL 200  
QY 54 RYDDG-YVSTSLRSALHAGOSILS-----GYSTYYIYVIATAPNMFVNDVLG---- 102  
Db 201 SKDNAYFYANYSGPLTYEDNELLSYFIEDIGNSYYIYFHNRFWFENGEGOLIGPLKE 260  
QY 103 -----VYSPHPYEQEVSALGGIPYSQIYGYWRVNFVGVIDERLHRRNEYDRYY 150  
Db 261 RGEIYVVYVYKILARYLERLANGLEIP-----RFNWL-----KYQTSYY 303  
QY 151 -----RNLNTAPADGYRLAGFPDPDQHWREEPWIHAPQCGGSSRTITGTCNEETQN 205  
Db 304 PLLSSYQLPFAQRNDYLA-----SGDNIN-DIQF 333  
QY 206 LSTIYLREYQSKVKROIIFSQYQSEVDIYN 234  
Db 334 IDT-YEKFTLQLQKQFKAYKQEVLDIYN 361  
RESULT 12  
S60935  
hypothetical protein YPL055c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein LPE17c  
C:Species: Saccharomyces cerevisiae  
C>Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 05-Nov-1999  
C:Accession: S60935  
R:Winnett, E.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Hall, J.; Storms, R.  
submitted to the EMBL Data Library, October 1995  
A:Reference number: S60921  
A:Molecule type: DNA  
A:Residues: 1-332 <WIN>  
A:Cross-references: EMBL:U39205; NID:g1079672; PID:g1079687; GSPDB:GN00016; MIPS:YPL0  
C:Genetics:  
A:Gene: MIPS:YPL055c  
A:Map position: 16L  
Query Match 6.8%; Score 88.5; DB 2; Length 332;  
Best Local Similarity 19.0%; Pred. No. 2.2;  
Matches 60; Conservative 38; Mismatches 94; Indels 123; Gaps 15;  
QY 2 GDRLYRADSRPPDE-----IKRS-----GGLMPRGHNEYEDRGTMNINLYDHARGTGTG 51  
Db 6 GNNYSYSTPPRGCGYHHARRSGGAGSYRGGNASY--GARYN--SDYEQPPQEG 60  
QY 52 FVR-----YDGGYVST--LSLSAHLAGO---SILSGYSTYYIYVIATAPNMFVNDVL 101  
Db 61 DLQGTGAYRNGVTDTRPYYSANSRHYQAQSPRYNNGTNSYHL-----PQRNSQDTN 114  
QY 102 G-----VYSPHPYEQEV-----SALGGIP-- 120  
Db 115 GRTTSAQEDNDEKRVKSRYNRQADHPROQPMVSGTSSRNGSSSSTSSSTSSNGLPPP 174  
QY 121 -----YSQIYGYWRVNFVGVIDERLHRRNEYDRYYRNLNIAPADGYRLAGFP 168

Db 175 PSVSSINRRSYHSSAVPY-----SSSHYNNYHHRETTPPPPPSGNGYAKGYP 222  
Qy 169 PDQAWREEPWIHAPOCGN---SSRITGDTCNEETQNLSTIYLREYQSKVK----- 219  
Db 223 V-----HVPENRNSDSSSSVYKKRILDMKDSFFIYLDKFNKAKTNTE 270  
Qy 220 -----RQIFSDYOS 228  
Db 271 SECEKAREVFKESDS 285

## RESULT 13

A25973

pertussis toxin chain S1 precursor - Bordetella bronchiseptica

C:Species: Bordetella bronchiseptica

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: A25973

R:Arico, B.; Rappuoli, R.

J. Bacteriol. 169, 2847-2853, 1987

A:Title: Bordetella parapertussis and Bordetella bronchiseptica contain transcriptional

A:Reference number: A25973; MUID:8722217

A:Accession: A25973

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-269 &lt;ARI&gt;

C:Superfamily: pertussis toxin chain S1

Query Match 6.7%; Score 88; DB 1; Length 269;  
Best Local Similarity 23.5%; Pred. No. 1.8;  
Matches 52; Conservative 23; Mismatches 62; Indels 84; Gaps 11;

Qy 5 LYRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNNINLYDHARGTQTGFVRYDDGYVSTSL 64

Db 41 VYRDSRPDEDVFON-GFTAWGND-----NVLEHLTGRSCQVGSNSAFVSTSS 89

Qy 65 SLR-----SAHLAQSIILSGYSTYIYVIATAPNMFNVNDVLGVYSPHYE 110

Db 90 SRRTEYVLEHRMQEAVEAERAGRG--TGHFYGIYIEV-RADNNF----- 131

Qy 111 QEVSALGGIPYSQIYGYRYVNFVIDERLHRNREYRDRYRNLM--IAPAEQGYRLAGFP 168

Db 132 -----YGAASSYF-----EYDVTYGDNAGRILAGA-----LATYQ 161

Qy 169 PDQAWREEPWIHAPOCGNSRPT-----ITGDFCNEETQN 205

Db 162 SEXLAHRRIP-----PENIRRVTRVYHNGITGETTTTEYPN 197

## RESULT 14

T46366

hypotheical protein DKFZp434C0118.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000

C:Accession: T46366

R:Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23031

A:Accession: T46366

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-547 &lt;AAA&gt;

A:Cross-references: EMBL:AL137520

A:Experimental source: adult testis; clone DKFZp434C0118

C:Genetics:

A:Note: DKFZp434C0118.1

Query Match 6.7%; Score 88; DB 2; Length 547;  
Best Local Similarity 23.4%; Pred. No. 4.7;  
Matches 39; Conservative 32; Mismatches 62; Indels 34; Gaps 9;

Qy 53 VRYDDGVVSTSLRSRAHLAQSIILSGY-----STYYIVVIATAPNMFNVNDVLGVYSP 106

Db 364 VRFDGVVISKTYIRQ-----GEQSLDGFYRAWHQVEYYRI-----RFFPDGHVWMLTTP 414  
Qy 107 HPYEQEVSAL--GGIPYSQIYGYRYVNF-----FGVIDERL-HRNREYRDRYRN 153  
Db 415 EEOOSIYPRLTRNTRTDAILLGHYRLSQDQDNQTKYFAVITKKKEKPLDYKYRYFRRV 474  
Qy 154 NIAPAEQGYR--LAGFPPDQAWREEPWIHAPOCGNSRRTITGDT 198  
Db 475 PVQEAQDSFHVGLQLCSSHQRFNKLIIWHHS---CHITYKS-TGET 517

## RESULT 15

T40574

guanine nucleotide binding protein beta subunit-like - fission yeast (Schizosaccharom

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T40574

R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: Z21938

A:Accession: T40574

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-809 &lt;SEE&gt;

A:Cross-references: EMBL:AL035226; PIDN:CAA22832.1; GSPDB:GN00067

A:Experimental source: strain 972h-; cosmid c609

C:Genetics:

A:Gene: SPAC609.03

A:Map position: 2

A:Introns: 21/1; 52/2; 280/3; 780/2

Query Match 6.7%; Score 87; DB 2; Length 809;  
Best Local Similarity 22.6%; Pred. No. 9.6;  
Matches 43; Conservative 29; Mismatches 54; Indels 64; Gaps 12;

Qy 1 NGD---RLYRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNNINLYDHARGTQTGFVRYDD 57

Db 110 SGDKLIKLFDLDS-----SKEGGM-----DHGMEQTQRCWSCALDSVKNIIVPCDN 154

Qy 58 GY-----VSTSLSLRSAHLAGO-----SILSGYSTYII--YVIATAPNMFNVNDVL 101

Db 155 GHTFLVCSEGTARQYDIRPHVCNQDLDCPSILVNTNPNRYNLYTITMSPS----- 206

Qy 102 GVYSPHPYEQEVSALGGI-PYSQIYGYRYVNFVIDERLHRNREYRDRYRNLIAPAE 160

Db 207 -----NPY---YFAIGGTHPYAFLY-----DRRMVK-KSFRDDW--TMNITSPEKD 245

Qy 161 GYRLAGFPPD 170

Db 246 CRCVRKFSPPD 255

Search completed: June 10, 2002, 17:32:32

Job time: 169 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:31:13 ; Search time 11.02 Seconds  
(without alignments)  
843.257 Million cell updates/sec

Title: US-09-297-171-1

Perfect score: 1304

Sequence: 1 NGRDLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNRIDEL 240

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1304	100.0	258	1 ELAP_ECOLI	P06717 escherichia
2	1292	99.1	258	1 ELAH_ECOLI	P43530 escherichia
3	1088	83.4	258	1 CHPA_VIBCH	P01555 vibrio chol
4	682	52.3	263	1 E2BA_ECOLI	P43528 escherichia
5	679.5	52.1	259	1 E2BA_ECOLI	P13810 escherichia
6	93	7.1	269	1 TOX1_BORPE	P04977 bordetella
7	91.5	7.0	471	1 CD36_RAT	Q07969 rattus norv
8	88.5	6.8	566	1 BGIC_MAIZE	P49235 zea mays (m
9	88	6.7	447	1 FBX9_HUMAN	Q9UK97 homo sapien
10	86	6.6	423	1 DHE3_AERPE	Q9YC65 aeropyrum p
11	84.5	6.5	826	1 CRAA_BACUH	Q9A597 bacillus th
12	83	6.4	254	1 PRIO_CRIGR	Q60506 cricetus
13	83	6.4	254	1 PRIO_CRIMI	Q60468 cricetus
14	82	6.3	256	1 PRP2_BOVIN	Q01880 bos taurus
15	82	6.3	257	1 PRIO_PIG	P49927 sus scrofa
16	82	6.3	446	1 SUN_SYNY3	P72943 synecocyst
17	82	6.3	470	1 F261_RAT	P07953 rattus norv
18	81.5	6.2	419	1 NSMA_MOUSE	O70572 mus musculu
19	81.5	6.2	828	1 MRKC_KLEPN	P21647 klebsiella
20	81	6.2	256	1 PRIO_CEREL	P79142 cervus elap
21	81	6.2	256	1 PRIO_FELCA	O18754 felis silve
22	81	6.2	264	1 PRIO_BOVIN	P10279 bos taurus
23	80.5	6.2	252	1 Y455_RICPR	Q9A288 rickettsia
24	80.5	6.2	880	1 LYTD_BACSU	P39648 bacillus su
25	80	6.1	254	1 PRIO_MESAU	P04273 mesocricetu
26	80	6.1	256	1 PRP2_TRAST	P40243 tragelaphus
27	80	6.1	264	1 PRP1_TRAST	P40242 tragelaphus
28	79.5	6.1	629	1 Z195_HUMAN	O14628 homo sapien
29	79	6.1	238	1 PRIO_THEGE	Q95270 theropithec
30	79	6.1	246	1 PRIO_CERTO	Q95176 cercocebus
31	79	6.1	260	1 PRIO_SAISS	P40258 saimir sci
32	79	6.1	474	1 CRT1_SYNP7	P26294 synecococc
33	78.5	6.0	255	1 PRIO_CAMDR	P79141 camelus dro

#### ALIGNMENTS

```
RESULT_1
ELAP_ECOLI
ID ELAP_ECOLI STANDARD; PRT; 258 AA.
AC P06717; P01554;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Heat-labile enterotoxin A chain precursor (LT-A, porcine) (LTP-A).
OS ELTA OR LTPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISOLATE PCG86, AND ISOLATE P307;
RX MEDLINE=87137303; PubMed=3546273;
RA Yamamoto T., Gojobori T., Yokota T.;
RT "Evolutionary origin of pathogenic determinants in enterotoxigenic
RT Escherichia coli and Vibrio cholerae O1.";
RL J. Bacteriol. 169:1352-1357(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ISOLATE P307;
RX Dykes C.W., Halliday I.J., Hobden A.N., Read M.J., Harford S.;
RT "A comparison of the nucleotide sequence of the A subunit of heat-
RT labile enterotoxin and cholera toxin.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ISOLATE P307;
RX MEDLINE=82167425; PubMed=6279611;
RA Spicer E.K., Noble J.A.;
RT "Escherichia coli heat-labile enterotoxin. Nucleotide sequence of the
RT A subunit gene.";
RL J. Biol. Chem. 257:5716-5721(1982).
RN [4]
RP SEQUENCE OF 19-258 FROM N.A.
RC STRAIN=ISOLATE P307;
RX MEDLINE=91093102; PubMed=2266142;
RA Tsuji T., Inoue T., Miyama A., Okamoto K., Honda T., Miwatani T.;
RT "A single amino acid substitution in the A subunit of Escherichia
RT coli enterotoxin results in a loss of its toxic activity.";
RN [5]
RP J. Biol. Chem. 265:22520-22525(1990).
RN [6]
RP SEQUENCE OF 1-40 FROM N.A.
RA Trachman J.D., Maas W.K.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RX MEDLINE=93240541; PubMed=8478941;
RA Sixma T.K., van Zanten B.A.M., Dauter Z., Hol W.G.J.;
RT "Refined structure of Escherichia coli heat-labile enterotoxin, a
RT close relative of cholera toxin.";
RL J. Mol. Biol. 230:890-918(1993).
RN [7]
```

34 78.5 6.0 1444 1 ADP1\_MYCGE P20796 mycoplasma  
35 78 6.0 252 1 PRIO\_ATEPA P51446 ateles pani  
36 78 6.0 256 1 PRIO\_CAPHI P52113 capra hircu  
37 78 6.0 256 1 PRIO\_ODOHE P47852 odocoileus  
38 78 6.0 257 1 PRIO\_MUSVI P40244 mustela vis  
39 77.5 5.9 173 1 Y028\_NPVAC P41432 autographa  
40 77 5.9 349 1 NTRB\_KLEPN P06218 klebsiella  
41 77 5.9 471 1 F261\_HUMAN P16118 homo sapien  
42 76.5 5.9 535 1 PYRG\_BACSU P13242 bacillus su  
43 76.5 5.9 556 1 SYQ\_VIBCH Q9Kta6 vibrio chol  
44 76.5 5.9 575 1 COE2\_MOUSE O08792 m transcrip  
45 76.5 5.9 596 1 MBHL\_RHILV P18636 rhizobium l

RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
RX MEDLINE=91238966; PubMed-2034287;  
RA Sixma T.K., Pronk S.E., Kalk K.H., Wartna E.S., van Zanten B.A.M.,  
RA Witholt B., Hol W.G.J.;  
RT "Crystal structure of a cholera toxin-related heat-labile enterotoxin  
RT from *E. coli*.";  
RL Nature 351:371-377(1991).  
RN [8]  
RP DISCUSSION OF SEQUENCE.  
RX MEDLINE=95349400; PubMed-7623669;  
RA Domenighini M., Pizsa M., Jobling M.G., Holmes R.K., Rappuoli R.;  
RT "Identification of errors among database sequence entries and  
RT comparison of correct amino acid sequences for the heat-labile  
RT enterotoxins of *Escherichia coli* and *Vibrio cholerae*.";  
RL Mol. Microbiol. 13:1165-1167(1995).  
CC -!- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY  
CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.  
CC -!- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M15361; AAA24791.1; -;  
DR EMBL; M15362; AAA24793.1; -;  
DR EMBL; M35581; AAA98202.1; -;  
DR EMBL; V00275; CAA23532.1; -;  
DR EMBL; M57244; AAB59161.1; -;  
DR EMBL; M61015; AAA24335.1; -;  
DR EMBL; A04913; CAA00402.1; -;  
DR PIR; A26946; QLECEA.  
DR PIR; A01817; QLECA.  
DR PDB; 1LTA; 31-JAN-94.  
DR PDB; 1LFB; 31-JUL-94.  
DR PDB; 1LFG; 15-SEP-95.  
DR PDB; 1LTI; 17-AUG-96.  
DR PDB; 1LTS; 31-JAN-94.  
DR PDB; 1LTT; 31-JAN-94.  
DR PDB; 1LTJ; 07-JUL-97.  
DR PDB; 1LR4; 16-JUN-97.  
DR PDB; 1LHTL; 20-APR-95.  
DR InterPro: IPR001144; Enterotoxin\_A.  
DR Pfam: PF01375; Enterotoxin\_A; 1.  
DR PRINTS: PR00771; ENTEROTOXINA.  
KW Enterotoxin; Signal; 3D-structure.  
FT SIGNAL 1 18  
FT CHAIN 19 258 HEAT-LABILE ENTEROTOXIN A CHAIN.  
FT DISULFID 205 217  
FT ACT\_SITE 130 130  
FT VARIANT 130 130  
FT CONFLICT 37 39  
FT CONFLICT 45 45  
FT CONFLICT 93 93  
FT CONFLICT 100 110  
FT CONFLICT 119 120  
FT CONFLICT 159 159  
FT CONFLICT 207 207  
FT STRAND 23 27  
FT HELIX 31 37  
FT TURN 38 38  
FT STRAND 39 40  
FT TURN 43 44  
FT TURN 48 49  
FT TURN 59 64  
FT HELIX 75 76  
FT STRAND 77 81  
FT HELIX 84 95  
FT STRAND 100 106  
FT STRAND 110 111

FT STRAND	112	114
FT HELIX	115	122
FT HELIX	126	128
FT STRAND	130	134
FT TURN	135	135
FT STRAND	137	138
FT HELIX	139	141
FT STRAND	142	149
FT TURN	150	151
FT STRAND	152	153
FT STRAND	158	159
FT TURN	161	162
FT HELIX	165	168
FT TURN	169	170
FT STRAND	174	174
FT HELIX	176	178
FT TURN	179	179
FT HELIX	180	182
FT TURN	187	188
FT HELIX	190	193
FT TURN	195	196
FT HELIX	197	200
FT TURN	203	204
FT HELIX	215	240
FT TURN	241	241
FT HELIX	242	244
FT HELIX	250	253
SQ SEQUENCE	258 AA; 29902 MW; 2F0786442619F81F CRC64;	

Query Match 100.0%; Score 1304; DB 1; Length 258;  
Best Local Similarity 100.0%; Pred. No. 1.5e-112;  
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	NGDRLYRADSRPPDEIKRSGGLMPRGHNEVDRGTOMNINLYDHARGTQTFVRYDDGYV	60
Db	19	NGDRLYRADSRPPDEIKRSGGLMPRGHNEVDRGTOMNINLYDHARGTQTFVRYDDGYV	78
QY	61	STSLSRSLAGQSILSGYSTYYIYVIATAPNMFNVNDVLGYSPHPYEQEVSALGGIP	120
Db	79	STSLSRSLAGQSILSGYSTYYIYVIATAPNMFNVNDVLGYSPHPYEQEVSALGGIP	138
QY	121	YSOIYGYRVNFCVIDERLHRNREYDRYRNLIAPAEDGYRLAGPPDPHQAWREPWI	180
Db	139	YSOIYGYRVNFCVIDERLHRNREYDRYRNLIAPAEDGYRLAGPPDPHQAWREPWI	198
QY	181	HHAPOCGNSSRTITGDTCTNEETQNLSTIYLYREVQSKVKRQIFSDYQSEVDIYNRIDEL	240
Db	199	HHAPOCGNSSRTITGDTCTNEETQNLSTIYLYREVQSKVKRQIFSDYQSEVDIYNRIDEL	258

RESULT 2  
ELAH\_ECOLI STANDARD; PRT; 258 AA.  
AC P43530;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Heat-labile enterotoxin A chain precursor (LT-A, human) (LTH-A).  
GN ELTA OR LTFA OR TOXA.  
OS *Escherichia coli*.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC *Escherichia*.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ISOLATE H10407;  
RC MEDLINE=84185610; PubMed-6325417;  
RA Yamamoto T., Tamura T., Yokota T.;  
RT "Primary structure of heat-labile enterotoxin produced by *Escherichia*  
RT *coli* pathogenic for humans.";  
RL J. Biol. Chem. 259:5037-5044(1984).  
RN [2]

```
RP REVISION TO 207.
RC STRAIN=ISOLATE H10407;
RX MEDLINE=87137303; PubMed=3546273;
RA Yamamoto T., Gojobori T., Yokota T.;
RT "Evolutionary origin of pathogenic determinants in enterotoxigenic
RL Escherichia coli and Vibrio cholerae O1.";
RN J. Bacteriol. 169:1352-1357(1987).
RC SEQUENCE FROM N.A.
RP STRAIN=ISOLATE H10407;
RX MEDLINE=9325225; PubMed=8486242;
RA Inoue T., Tsuji T., Koto M., Imamura S., Miyama A.;
RT "Amino acid sequence of heat-labile enterotoxin from chicken
RL H 10407.";
RN FEMS Microbiol. Lett. 108:157-161(1993).
RC DISCUSSION OF SEQUENCE.
RP MEDLINE=95349400; PubMed=7623669;
RA Domenighini M., Pizzi M., Jobling M.G., Holmes R.K., Rappuoli R.;
RT "Identification of errors among database sequence entries and
RL comparison of correct amino acid sequences for the heat-labile
RC enterotoxins of Escherichia coli and Vibrio cholerae.";
RX Mol. Microbiol. 15:1165-1167(1995).
CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY
CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.
CC -1- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01995; AAA24685.1; -.
DR EMBL; S60731; CAB31891.1; -.
DR HSPF; P06717; 1LTG.
DR InterPro; IPR001144; Enterotoxin_A.
DR Pfam; PF01375; Enterotoxin_A; 1.
DR PRINTS; PR00771; ENTEROTOXINA.
KW Enterotoxin; Signal.
FT SIGNAL 1 18 BY SIMILARITY.
FT CHAIN 19 258 HEAT-LABILE ENTEROTOXIN A CHAIN.
FT DISULFID 205 217 BY SIMILARITY.
FT ACT_SITE 128 128 BY SIMILARITY.
FT CONFLICT 207 207 N -> D (IN REF. 1 AND 3).
SQ SEQUENCE 258 AA; 29872 MW; 2BB153C777FD78B7 CRC64;

Query Match 99.1%; Score 1292; DB 1; Length 258;
Best Local Similarity 98.8%; Pred. No. 1.9e-111;
Matches 237; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NGDRLYRADSRPPDIKRSGLMPRHNEYFDRGQTOMNLINLDHARGTQTVGRVYDDGVV 60
Db 19 NGDRLYRADSRPPDIKRSGLMPRHNEYFDRGQTOMNLINLDHARGTQTVGRVYDDGVV 78
Qy 61 STSLSRSAHLAQSIKSGSYTYIYVATAPNMVNDVLGVSPHPYQEVSAALGGIP 120
Db 79 STSLSRSAHLAQSIKSGSYTYIYVATAPNMVNDVLGVSPHPYQEVSAALGGIP 138
Qy 121 YSQIYGVYRVNFGVDERLHRRNREYDRYRNLTAPADGDYVLGFPDPHQAREEPMI 180
Db 139 YSQIYGVYRVNFGVDERLHRRNREYDRYRNLTAPADGDYVLGFPDPHQAREEPMI 198
Qy 181 HHAPOGCGNSSRTITGDTCTNEETQNLSTLYLREYQSKVKRQIFSDYQSEVDIYNIRDEL 240
Db 199 HHAPOGCGNSSRTITGDTCTNEETQNLSTLYLREYQSKVKRQIFSDYQSEVDIYNIRDEL 258

RESULT 3
```

```
CHTA_VIBCH STANDARD; PRT; 258 AA.
AC P01555; Q56634; Q9JVPV1;
DT 21-JUL-1986 (Rel. 01, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cholera enterotoxin, A chain precursor (NAD(+)-diphthamide ADP-
DE ribosyltransferase) (EC 2.4.2.36) (Cholera enterotoxin A subunit).
GN CXTA OR TOXA OR VC1457.
OS Vibrio cholerae.
OX Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN NCBI_TaxID=666;
RC SEQUENCE FROM N.A.
RP STRAIN=EL TOR 2125;
RX MEDLINE=84068199; PubMed=6646234;
RA Mekalanos J.J., Swartz D.J., Pearson G.D.N., Harford N., Groyne F.,
RA de Wilde M.;
RT "Cholera toxin genes: nucleotide sequence, deletion analysis and
RL vaccine development.";
RN Nature 306:551-557(1983).
RC SEQUENCE FROM N.A.
RP STRAIN=CLASSICAL 569B / ATCC 25870 / SEROTYPE O1;
RX MEDLINE=91355224; PubMed=1883840;
RA Dams E., de Wolf M., Dierick W.;
RT "Nucleotide sequence analysis of the CT operon of the Vibrio cholerae
RL classical strain 569B.";
RN Biochim. Biophys. Acta 1090:139-141(1991).
RC SEQUENCE FROM N.A.
RP STRAIN=1854 / O139-BENGAL;
RX Yamamoto K., Do V.G.R.F., Xu M., Iida T., Miwatani T., Albert M.J.,
RA Honda T.;
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
RC SEQUENCE FROM N.A.
RP STRAIN=EL TOR 2125;
RX Dams E., de Wolf M., Dierick W.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
RC SEQUENCE FROM N.A.
RP STRAIN=KNIH002;
RA Shin H.J., Park Y.C., Kim Y.C.;
RT "Cloning and nucleotide sequence analysis of the virulence gene
RL cassette from Vibrio cholerae KNIH002 isolated in Korea.";
RN Misaimurhag Hoiji 35:205-210(1999).
RC SEQUENCE FROM N.A.
RP STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RL cholerae.";
RN Nature 406:477-483(2000).
RC SEQUENCE OF 1-212 FROM N.A.
RP STRAIN=CLASSICAL 569B / ATCC 25870 / SEROTYPE O1;
RX MEDLINE=85006737; PubMed=6090390;
RA Lockman H.A., Galen J.E., Kaper J.B.;
RT "Vibrio cholerae enterotoxin genes: nucleotide sequence analysis of
RL DNA encoding ADP-ribosyltransferase.";
RN J. Bacteriol. 159:1086-1089(1984).
RC SEQUENCE OF 213-258 FROM N.A.
RP MEDLINE=84061784; PubMed=6315707;
RA Lockman H., Kaper J.B.;
```

RT "Nucleotide sequence analysis of the A2 and B subunits of Vibrio  
 RT cholerae enterotoxin.",  
 RL J. Biol. Chem. 258:13722-13726(1983).  
 RN [9]  
 RP SEQUENCE OF 19-27.  
 RX MEDLINE=81212799; PubMed=7238869;  
 RA Duffy L.K., Peterson J.W., Kurosky A.;  
 RT "Isolation and characterization of a precursor form of the 'A'  
 RT subunit of cholera toxin.",  
 RL FEBS Lett. 126:187-190(1981).  
 RN [10]  
 RP SEQUENCE OF 19-38 AND 213-232.  
 RX MEDLINE=76259136; PubMed=955672;  
 RA Klapper D.G., Finkelstein R.A., Capra J.D.;  
 RT "Subunit structure and N-terminal amino acid sequence of the three  
 RT chains of cholera enterotoxin.",  
 RL Immunochimistry 13:605-611(1976).  
 RN [11]  
 RP SEQUENCE OF 27-72 AND 111-139.  
 RX MEDLINE=79169830; PubMed=437113;  
 RA Lai C.-Y., Cancedda F., Chang D.;  
 RT "Primary structure of cholera toxin subunit A1: isolation, partial  
 RT sequences and alignment of the BrCN fragments.",  
 RL FEBS Lett. 100:85-89(1979).  
 RN [12]  
 RP SEQUENCE OF 213-258.  
 RX MEDLINE=82053094; PubMed=7028752;  
 RA Duffy L.K., Peterson J.W., Kurosky A.;  
 RT "Covalent structure of the gamma chain of the A subunit of cholera  
 RT toxin.",  
 RL J. Biol. Chem. 256:12252-12256(1981).  
 RN [13]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RX MEDLINE=95387395; PubMed=7658473;  
 RA Zhang R.G., Scott D.L., Westbrook M.L., Nance S., Spangler B.D.,  
 RA Shipley G.G., Westbrook E.M.;  
 RT "The three-dimensional crystal structure of cholera toxin.",  
 RL J. Mol. Biol. 251:563-573(1995).  
 CC -!- FUNCTION: THE ALPHA/GAMMA CHAIN (A SUBUNIT) IS AN ADP-RIBOSYLATING  
 CC TOXIN.  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + peptide diphthamide = nicotinamide +  
 CC peptide N-(ADP-D-ribosyl)diphthamide.  
 CC -!- SUBUNIT: CONTAINS 3 KINDS OF CHAINS, AN ALPHA AND A GAMMA CHAIN  
 CC (FROM THE SAME PRECURSOR MOLECULE), LINKED BY AN INTERCHAIN  
 CC DISULFIDE BOND, ASSOCIATE NONCOVALENTLY WITH AN AGGREGATE OF 4 TO  
 CC 6 BETA CHAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X00171; CAA24995.1; -  
 DR EMBL; X58785; CAA1590.1; -  
 DR EMBL; D30053; BAA06290.1; -  
 DR EMBL; X58786; CAA41592.1; -  
 DR EMBL; K02679; AAA27514.1; -  
 DR EMBL; AF175708; AAD51359.1; -  
 DR EMBL; AE004224; AAF94614.1; -  
 DR EMBL; K01170; AAA27572.1; -  
 DR EMBL; D30052; BAA06288.1; -  
 DR PIR; A01818; XVCA. -  
 DR PIR; A05129; A05129. -  
 DR PIR; S14625; S14625. -  
 DR PIR; S17665; S17665. -  
 DR PIR; S14623; S14623. -  
 DR PDB; 1XPB; 01-APR-97. -  
 DR PDB; 1XTC; 01-AUG-96. -  
 DR TIGR; VC1457; -  
 DR InterPro; IPR001144; Enterotoxin\_A.

DR Pfam; PF01375; Enterotoxin\_A; 1.  
 DR PRINTS; PR00771; ENTEROTOXIN.  
 KW Enterotoxin; Signal; NAD; Transferase; Glycosyltransferase;  
 KW 3D-structure; Complete proteome.  
 FT SIGNAL 1 18  
 FT CHAIN 19 212 CHOLERA ENTEROTOXIN, CHAIN-A1 (ALPHA).  
 FT CHAIN 213 258 CHOLERA ENTEROTOXIN, CHAIN-A2 (GAMMA).  
 FT DISULFID 217 217 INTERCHAIN (WITH GAMMA CHAIN).  
 FT ACT\_SITE 62 62 INTERACT WITH NAD (BY SIMILARITY).  
 FT ACT\_SITE 130 130 BY SIMILARITY.  
 FT CONFLICT 20 20 D -> N (IN REF. 9).  
 FT CONFLICT 37 37 S -> R (IN REF. 10).  
 FT CONFLICT 39 39 G -> L (IN REF. 11).  
 FT CONFLICT 45 46 QS -> SE (IN REF. 11).  
 FT CONFLICT 111 111 N -> L (IN REF. 11).  
 FT CONFLICT 132 132 S -> A (IN REF. 11).  
 FT CONFLICT 213 213 M -> I (IN REF. 1).  
 FT CONFLICT 247 248 DI -> ID (IN REF. 12).  
 FT CONFLICT 256 256 D -> N (IN REF. 12).  
 SQ SEQUENCE 258 AA; 29335 MW; 0F7EBAE62069A5D0 CRC64;

Query Match 83.4%; Score 1088; DB 1; Length 258;  
 Best Local Similarity 81.7%; Pred. No. 1e-92;

Matches 196; Conservative 24; Mismatches 20; Indels 0; Gaps 0;

QY 1 NGRDLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGTVRYDDGYV 60  
 Db 19 NDKLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGTVRYDDGYV 78  
 QY 61 STSLSLSAHLAQSILSGSYTYIYVIATAPNMFNNDVLGVSPHPYEQEVSALGGIP 120  
 Db 79 STSLSLSAHLVGTILSGHSTYIYVIATAPNMFNNDVLGVSPHPYEQEVSALGGIP 138  
 QY 121 YSQTGYHYRVNFGVIDERLHNRNREYRDYRNLNIAEADGRLAGPPDQHQAREPWI 180  
 Db 139 YSQTGYHYRVHFGVLDQLHNRNREYRDYRNLNIAEADGRLAGPPDQHQAREPWI 198  
 QY 181 HHAPOGCGNSRTTGTCTCNEETONLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIDEL 240  
 Db 199 HHAPPGCGNAPRSMSTNCTDEKTSGLGVKELDEYQSKVKRQIFSGYQSDIDTHNRKDEL 258

RESULT 4  
 E2BA\_ECOLI STANDARD; PRT; 263 AA.  
 AC P43528;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Heat-labile enterotoxin IIB, A chain precursor (LT-IIB).  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=ISOLATE 41.  
 RA MEDLINE=89359131; PubMed=2670900;  
 RA Pickett C.L., Twiddy E.M., Coker C., Holmes R.K.;  
 RT "Cloning, nucleotide sequence, and hybridization studies of the type  
 RT IIB heat-labile enterotoxin gene of Escherichia coli.",  
 RL J. Bacteriol. 171:4945-4952(1989).  
 RN [2]  
 RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).  
 RX MEDLINE=96399713; PubMed=8805549;  
 RA van den Akker F., Sarfaty S., Twiddy E.M., Connell T.D., Holmes R.K.,  
 RA Hol W.G.;  
 RT "Crystal structure of a new heat-labile enterotoxin, LT-IIB.",  
 RL Structure 4:665-678(1996).  
 CC -!- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY  
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.  
 CC -!- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M28523; AAA53285.1; -;  
 CC PDB; 1TII; PRELIMINARY.  
 CC InterPro; IPR001144; Enterotoxin A.  
 CC Pfam; PF01375; Enterotoxin A; 1.  
 CC PRINTS; PR00771; ENTEROTOXINA.  
 CC Enterotoxin; Signal; 3D-structure.  
 CC SIGNAL 1 20  
 CC FT CHAIN 21 263 HEAT-LABILE ENTEROTOXIN IIB, A CHAIN.  
 CC FT DISULFID 205 217  
 CC FT ACT\_SITE 130 130  
 CC SEQUENCE 263 AA; 29485 MW; A56E8A069427CFB9 CRC64;  
 CC -----  
 CC Query Match 52.3%; Score 682; DB 1; Length 263;  
 CC Best Local Similarity: 57.4%; Pred. No. 1.9e-55;  
 CC Matches 124; Conservative 35; Mismatches 57; Indels 0; Gaps 0;  
 CC  
 CC QY 6 YRDSRPDEIKRSGGLMPRGHNEYFDRCGTOMINLYDHARGTGTGFRVYDDGVVSTLS 65  
 CC DB 24 FRASRTPEVRSRGLIPRGQDEAYERGTPININLYDHARGTGTGFRVYDDGVVSTTT 83  
 CC QY 66 LRSALHAGQSGILSGSYTYIYVATAPNMVNDVGVSPHPYEQVQSALGGIPYSQIY 125  
 CC DB 84 LRAHLLGQNLGNGYNYIYVVAAPNLFVNGVGLGRYSYPSENEFAALGGIPLSQII 143  
 CC QY 126 GWYRVNFGVIDERLHNRNRYDRYRLNTAPADGSLAGFPDPHQAWREEPWIHAPQ 185  
 CC DB 144 GWYRVSGAEGGMHRRNRYDRGLFGLSAAPNEDGVRIAGFPDPGFAWEVPMREAPN 203  
 CC QY 186 CGNSRTITGDCNTEETQNLSTIYLYREYQSKVKRQ 221  
 CC DB 204 SCLPNKASDDTTCASLTNKLQHDADFKKYIKR 239  
 CC  
 CC RESULT 5  
 CC E2AA\_ECOLI STANDARD; PRT; 259 AA.  
 CC ID E2AA\_ECOLI  
 CC AC FL3810;  
 CC DT 01-JAN-1990 (Rel. 13, Created)  
 CC DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 CC DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 CC DE Heat-labile enterotoxin IIA, A chain precursor (LT-IIA).  
 CC OS Escherichia coli.  
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC OC Escherichia.  
 CC OC NCBI\_TaxID=562;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=88032841; PubMed=2822667;  
 CC RA Pickett C.L., Weinstein D.L., Holmes R.K.;  
 CC RT "Genetics of type IIA heat-labile enterotoxin of Escherichia coli:  
 CC RT operon fusions, nucleotide sequence, and hybridization studies.";  
 CC RL J. Bacteriol. 169:5180-5187(1987).  
 CC CC -1- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY  
 CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.  
 CC CC -1- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M17894; AAA24093.1; -;  
 CC DR PIR; A29831; A29831.  
 CC DR HSP; P43528; ITII.  
 CC InterPro; IPR001144; Enterotoxin A.  
 CC Pfam; PF01375; Enterotoxin A; 1.  
 CC PRINTS; PR00771; ENTEROTOXINA.  
 CC Enterotoxin; Signal; 18  
 CC FT SIGNAL 1 18  
 CC FT CHAIN 19 259 HEAT-LABILE ENTEROTOXIN IIA, A CHAIN.  
 CC FT DISULFID 203 215 BY SIMILARITY.  
 CC FT ACT\_SITE 128 128 BY SIMILARITY.  
 CC SEQUENCE 259 AA; 29242 MW; 996F311A32CABEA CRC64;  
 CC -----  
 CC Query Match 52.1%; Score 679.5; DB 1; Length 259;  
 CC Best Local Similarity 54.6%; Pred. No. 3.2e-55;  
 CC Matches 130; Conservative 39; Mismatches 66; Indels 3; Gaps 2;  
 CC  
 CC QY 6 YRDSRPDEIKRSGGLMPRGHNEYFDRCGTOMINLYDHARGTGTGFRVYDDGVVSTLS 65  
 CC DB 22 FRASRTPEVRSRGLIPRGQDEAYERGTPININLYDHARGTGTGFRVYDDGVVSTVT 81  
 CC QY 66 LRSALHAGQSGILSGSYTYIYVATAPNMVNDVGVSPHPYEQVQSALGGIPYSQIY 125  
 CC DB 82 LRAHLLGQNLGNGYNYIYVVAAPNLFVNGVGLGRYSYPSENEFAALGGIPLSQII 141  
 CC QY 126 GWYRVNFGVIDERLHNRNRYDRYRLNTAPADGSLAGFPDPHQAWREEPWIHAPQ 185  
 CC DB 142 GWYRVSGAEGGMHRRNRYDRGLFGLTAPNEDGVRIAGFPDPGFAWEVPMREAPN 201  
 CC QY 186 CGNSRTITGDCNTEETQNLSTIYLYREYQSKVKRQI-FSDYQSEVDI--YNRIRDEL 240  
 CC DB 202 QCVPNKKEFGKGVICISATNVLKSLYDLNFKLLKRLALTFMSEDDDFGVHGERDEL 259  
 CC  
 CC RESULT 6  
 CC TOXI\_BORPE STANDARD; PRT; 269 AA.  
 CC ID TOXI\_BORPE  
 CC AC P04977;  
 CC DT 13-AUG-1987 (Rel. 05, Created)  
 CC DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 CC DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 CC DE Pertussis toxin subunit 1 precursor (PTX S1) (Islet-activating protein  
 CC DE S1) (IAP S1) (NAD-dependent ADP-ribosyltransferase (EC 2.4.2.-)).  
 CC GN PTXA.  
 CC OS Bordetella pertussis.  
 CC OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
 CC OC Bordetella.  
 CC OC NCBI\_TaxID=520;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=BP165;  
 CC RX MEDLINE=86259651; PubMed=2873570;  
 CC RA Nicotia A., Perugini M., Franzini C., Casagli M.C., Borri M.G.,  
 CC RA Antoni G., Almonni M., Neri P., Ratti G., Rappuoli R.;  
 CC RT "Cloning and sequencing of the pertussis toxin genes: operon  
 CC RT structure and gene duplication.";  
 CC RL Proc. Natl. Acad. Sci. U.S.A. 83:4631-4635(1986).  
 CC RN [2]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=10536;  
 CC RX MEDLINE=90045957; PubMed=2554254;  
 CC RA Loomore S.M., Cunningham J.D., Bradley W.R., Yao E.L., Dekaban G.A.,  
 CC RA Klein M.H.;  
 CC RT "A unique sequence of the Bordetella pertussis toxin operon.";  
 CC RL Nucleic Acids Res. 17:8365-8365(1989).  
 CC RN [3]  
 CC RP SEQUENCE FROM N.A.  
 CC RX MEDLINE=86208173; PubMed=3704651;  
 CC RA Loch C., Keith J.M.;  
 CC RT "Pertussis toxin gene: nucleotide sequence and genetic organization.";  
 CC Science 232:1258-1264(1986).





```

Db 432 WYMPGLKLLIMKKNYGNPIYITENGIGDVTKETPLPMEALNDYKRLDYIQRH 491
QY 222 IFSDYQSEVDIYNRIR 237
Db 492 I-ATLKESIDLGSNVQ 506

RESULT 9
FBX9_HUMAN
ID FBX9_HUMAN STANDARD; PRT; 447 AA.
AC Q9UK37; O75986;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE F-box only protein 9.
GN FBX09 OR FBX9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20003061; PubMed-10531037;
RA Winston J.T., Koepf D.M., Zhu C., Ellledge S.J., Harper J.W.;
RT "A family of mammalian F-box proteins."
RL Curr. Biol. 9:1180-1182(1999).
RN [2]
RP SEQUENCE OF 121-447 FROM N.A.
RX MEDLINE-20003060; PubMed-10531035;
RA Cenciarrelli C., Chlaud D.S., Guardavaccaro D., Parks W., Vidal M.,
RA Pagano M.;
RT "Identification of a family of human F-box proteins."
RL Curr. Biol. 9:1177-1179(1999).
RN [3]
RP SEQUENCE OF 121-447 FROM N.A.
RA Tubby B.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX
CC (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 TPR REPEAT.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collabor
CC between the Swiss Institute of Bioinformatics and the EMBL outstat
CC The European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no
CC modified and this statement is not removed. Usage by and for comm
CC entities requires a license agreement (See http://www.isb-sib.ch/anno
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF176704; AAF03704.1; -
CC EMBL; AF174597; AAF04518.1; -
CC EMBL; AL031178; CAA02122.1; -
CC InterPro; IPR001810; F-box.
CC InterPro; IPR001440; TPR.
CC Pfam; PF00646; F-box; 1.
CC Pfam; PF00515; TPR; 1.
CC PROSITE; PS50181; FBOX; 1.
CC Ubiquitin conjugation; TPR repeat.
CC REPEAT 94 127
CC TPR.
CC DOMAIN 185 236
CC F-BOX.
CC SEQUENCE 447 AA; 52329 MW; 2A88163DAB889D69 CRC64;

Query Match 6.7%; Score 88; DB 1; Length 447;
Best Local Similarity 23.4%; Pred. No. 1.3;
Matches 39; Conservative 32; Mismatches 62; Indels 34; Gaps
QY 53 VRYDDGVSTSLRSALHAGQSILGSY-----STYYIYVIATAPNMFNVNDVLGVSP 106

```

```
Db 264 VRFDGVYISKTYIRO----GEOSLDGFYRAWHQVEYRYI-----RFPDGHVWMLTTP 314
QY 107 HPEYQEVSA-----GGIPYQIYGWYRW-----FGVIDERL-HRNREYDRYRNL 153
Db 315 EEPQSIYVPLRTNRTNRTDAILLGHYLSQDTONQTVFAVITKKKEKPLDYKYRFRV 374
QY 154 NIAPADGYR--LAGFPDQAWREEPWIHAPOGCGNSRTITGDT 198
Db 375 PVOEADQSFHVLQLCSSGHRFNKLWIHHS---CHITYKS-TGET 417

RESULT 10
DHE3_AERPE
ID DHE3_AERPE STANDARD; PRT; 423 AA.
AC Q9YC65;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
GN GDHA OR AP1386.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoynaga A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo H. Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -L-CATALYTIC ACTIVITY: L-glutamate + H(2)O + NAD(P)(+) = 2-
CC oxoglutarate + NH(3) + NAD(P)H.
CC -L- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -L- SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP000061; BNA80383.1; -
CC HSP; P96110; I826.
CC InterPro; IPR002086; Aldehyde_dehydr.
CC InterPro; IPR001625; GLFV_dehydrog.
CC InterPro; IPR00205; NAD_binding.
CC Pfam; PF0208; GLFV_dehydrog; 1.
CC Pfam; PF02812; GLFV_dehydrog; 1.
CC PRINTS; PR00082; GLFDHNRGNASE.
CC PROSITE; PS00074; GLFV_DEHYDROGENASE; 1.
KW Oxidoreductase; NAD; NADP; Complete proteome.
FT ACT_SITE 110 110 BY SIMILARITY.
FT NP_BIND 222 228 NAD (POTENTIAL).
SQ SEQUENCE 423 AA; 46697 MW; A4B0EAL07A89DA62 CRC64;

Query Match 6.6%; Score 86; DB 1; Length 423;
Best Local Similarity 21.9%; Pred. No. 1.8;
Matches 41; Conservative 32; Mismatches 58; Indels 56; Gaps 10;

QY 11 RPDETKRSGGLMPR-----GHNEYF-----DRGTQMNNINYDHARGTQTFVRYDDG 58
Db 7 OPTDPLEEARAQLRRRAVDLLGYDDYVVEVLNPRDLVRVT-----IKMDG 54
```

```
QY 59 YVSTSLRSALAGOSILSGYSTYYIYVIATAPNMFNVNDVLCVSPHPYEQVVSALGG 118
Db 55 TVTFLGWSQHSALGPKYGGVRYH-----PNV-TMNEVIALSMMTWK-----NSLAG 103
QY 119 IPYSQIYGWYRVNFGVIDER---LHRNREYDRYRNLN-----IAPADGYRLAGF 167
Db 104 LPYGGGKGVRVNPVKILSPRELELSR-----KYFESIDIVGVQDQIPAPVYT---- 153
QY 168 PPDQHW 174
Db 154 DPQVMSW 160

RESULT 11
CRAA_BACUH
ID CRAA_BACUH STANDARD; PRT; 826 AA.
AC Q9S597;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Pesticidal crystal protein cry27Aa (insecticidal delta-endotoxin
DE CryXXVIIA(a)) (Crystalline entomocidal protoxin) (94 kDa crystal
DE protein).
GN CRY27AA OR CRYXXVIIA(A).
OS Bacillus thuringiensis (subsp. higo).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=132266;
RN [1]
RP SEQUENCE FROM N.A.
RA Saitoh H.;
RT "94kDa mosquitocidal toxin from serovar higo.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
CC -L- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
CC EPITHELIAL CELLS OF INSECTS.
CC -L- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
CC OF THE SPORE COAT.
CC -L- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
CC N-TERMINUS.
CC -L- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB023293; BAA82796.1; -
CC InterPro; IPR001178; Endotoxin.
CC Pfam; PF00555; endotoxin; 1.
KW Toxin; Sporulation.
SQ SEQUENCE 826 AA; 94434 MW; DFD5313C5B6023E2 CRC64;

Query Match 6.5%; Score 84.5; DB 1; Length 826;
Best Local Similarity 21.3%; Pred. No. 5.9;
Matches 44; Conservative 39; Mismatches 83; Indels 41; Gaps 10;

QY 42 YDHARGTQTFVRYDDGVVSTSLSLSAHLAGOSILSGYSTYYIYVIATAPNMFNVN-DV 100
Db 626 YFRARHEHFYIEFD-----TTFSLRNSGQLEHLL-----HIY----PNTTKISGQ 670
QY 101 LGVYSPHPYEQVVSALGGIPYSQI---YGWYRVNFGVIDERLHRNREYDRYRNLNIAIP 157
Db 671 LLIIDKITEFPV-----GIFLQNTSEGYNTYDQNTSYNQNNYNNQNMDDTYQ-----P 720
QY 158 AEDGYRLAGFPDQAWREEPWIHAPOGCGNSRSTITGDTCTNLETQNLSTIYIUREYOS- 216
Db 721 NYDNYK-----QNSSGMYDNPY-NONPKDSYNQNTYTDYSDSYNQSNGVSGNYNQYNTY 774
```



```

CC PROPEP 232 254 REMOVED IN MATURE FORM.
CC LIPID 231 231 GPI-ANCHOR.
CC DOMAIN 90 231 PRP27-30 (PROTEASE RESISTANT CORE).
CC CARBOHYD 181 181 N-LINKED (GLCNAC. .) (BY SIMILARITY).
CC FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (BY SIMILARITY).
CC FT DISULFID 179 214 BY SIMILARITY.
CC FT DOMAIN 51 91 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
CC Q.
CC REPEAT 51 59 1.
CC REPEAT 60 67 2.
CC REPEAT 68 75 3.
CC REPEAT 76 83 4.
CC REPEAT 84 91 5.
CC SEQUENCE 254 AA; 27855 MW; 78963FC6F779D0F CRC64;

Query Match 6.4%; Score 83; DB 1; Length 254;
Best Local Similarity 25.0%; Pred. No. 1.8;
Matches 31; Conservative 20; Mismatches 51; Indels 22; Gaps 6;

QY 114 SALGGIPYSQIYGWYRVNFGVIDERLHNRREYRDYR- NLNIAPAEDGYRLAGFPDQH 172
DB 120 AVVGL- - - - -GGYMLGSAMSRPLHFGNDWEDRYRENMRYPNQVYR- - - - -PVDQ 168
QY 173 AWREBPWIHHAPOGCGN- - - - -SSRTITGDCNEETQNLSTIYLREYQSKVKRQIFSDYQSE 229
DB 169 YNQNNFVH- - - - -DCVNITIKQHTVTTTKGE- - - - -NFTETDKMMERVVEQMCVTQYQKE 221
QY 230 VDIY 233
DB 222 SQAY 225
```

```

RESULT 14
PRP2_BOVIN STANDARD; PRT; 256 AA.
AC Q01880;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Major prion protein 2 precursor (PrP) (Major scrapie-associated fibril
DE protein 2).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93118243; PubMed=1362024;
RA Yoshimoto J., Ihnuma T., Ishiguro N., Horiuchi M., Imamura M.,
RA Shinagawa M.;
RT "Comparative sequence analysis and expression of bovine PrP gene in
RT mouse L-929 cells.";
RL Virus Genes 6:343-356(1992).
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND ANIMALS
CC INFECTED WITH DEGENERATIVE NEUROLOGICAL DISEASES SUCH AS KURU,
CC CRUFTFELDT-JAKOB DISEASE (CJD), GERSTMANN-STRAUSSLER SYNDROME
CC (GSS), SCRAPIE, BOVINE SPONGIFORM ENCEPHALOPATHY (BSE),
CC TRANSMISSIBLE MINK ENCEPHALOPATHY (TME), ETC.
CC -!- SIMILARITY: BELONGS TO THE PRION FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D10614; BAA01469.1; -.
CC PIR; J00268; J00268.
CC HSSP; P04156; IEIG.
CC InterPro; IPR001610; PAC.
CC InterPro; IPR000817; Prion.
CC Pfam; PF00377; prion; 1.
CC PRINTS; PR00341; PRION.
CC SMART; SM00086; PAC; 1.
CC SMART; SM00157; PRP; 1.
CC PROSITE; PS00291; PRION_1; 1.
CC PROSITE; PS00706; PRION_2; 1.
KW Prion; Brain; Glycoprotein; GPI-anchor; Repeat; Signal.
FT SIGNAL 1 24
FT CHAIN 25 256 MAJOR PRION PROTEIN 2.
FT CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 182 217 BY SIMILARITY.
FT DOMAIN 54 95 5 X 8 AA TANDEM REPEATS OF P-H-G-G-G-W-G-
FT Q.
FT REPEAT 54 62 1.
FT REPEAT 63 70 2.
FT REPEAT 71 78 3.
FT REPEAT 79 86 4.
FT REPEAT 87 95 5.
FT SEQUENCE 256 AA; 27880 MW; 0D969FF2D9033B30 CRC64;

Query Match 6.3%; Score 82; DB 1; Length 256;
Best Local Similarity 24.2%; Pred. No. 2.2;
Matches 30; Conservative 21; Mismatches 51; Indels 22; Gaps 6;

QY 114 SALGGIPYSQIYGWYRVNFGVIDERLHNRREYRDYR- NLNIAPAEDGYRLAGFPDQH 172
DB 123 AVVGL- - - - -GGYMLGSAMSRPLHFGNDYEDRYRENMRYPNQVYR- - - - -PVDQ 171
QY 173 AWREBPWIHHAPOGCGN- - - - -SSRTITGDCNEETQNLSTIYLREYQSKVKRQIFSDYQSE 229
DB 172 YNQNNFVH- - - - -DCVNITVKEHTVTTTKGE- - - - -NFTETDKMMERVVEQMCITQYQRE 224
QY 230 VDIY 233
DB 225 SQAY 228
```

```

RESULT 15
PRIO_PIG STANDARD; PRT; 257 AA.
AC P49927;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major prion protein precursor (PrP).
GN PRNP.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95244627; PubMed=7727546;
RA Martin T., Hughes S., Hughes K., Dawson M.;
RT "Direct sequencing of PCR amplified pig PrP genes.";
RL Biochim. Biophys. Acta 1270:211-214(1995).
CC -!- FUNCTION: THE FUNCTION OF PRP IS NOT KNOWN. PRP IS ENCODED IN THE
CC HOST GENOME AND IS EXPRESSED BOTH IN NORMAL AND INFECTED CELLS.
CC -!- SUBUNIT: PRP HAS A TENDENCY TO AGGREGATE YIELDING POLYMERS CALLED
CC "RODS".
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- DISEASE: FOUND IN HIGH QUANTITY IN THE BRAIN OF HUMANS AND ANIMALS
CC INFECTED WITH DEGENERATIVE NEUROLOGICAL DISEASES SUCH AS KURU,
```





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:30:53 ; Search time 26.47 Seconds  
(without alignments)  
1568.522 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 1304  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNRDEL 240

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriopl:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1290	98.9	258	2	066280
2	357	27.4	68	2	047182
3	152	11.7	46	2	057372
4	94	7.2	850	5	09U804
5	93	7.1	269	2	069258
6	93	7.1	850	5	09GV36
7	91	7.0	392	10	081497
8	90.5	6.9	702	5	024995
9	89	6.8	269	2	093V22
10	88.5	6.8	332	3	002796
11	88	6.7	269	2	070057
12	87	6.7	547	4	09NT57
13	87	6.7	809	3	094527
14	86.5	6.6	145	16	09JZ10
15	85.5	6.6	256	6	095N12
16	85.5	6.6	435	17	09YB04

17	84	6.4	1219	12	09IMU3
18	83.5	6.4	524	5	09VLE2
19	83.5	6.4	715	5	0961P3
20	83.5	6.4	870	2	003988
21	83	6.4	353	16	097N09
22	83	6.4	7829	5	018559
23	82.5	6.3	888	3	09UVU7
24	82.5	6.3	215	6	097904
25	82	6.3	216	6	09TV00
26	82	6.3	235	6	097695
27	82	6.3	446	5	09WOC6
28	82	6.3	556	3	096UL8
29	81.5	6.2	419	11	070572
30	81.5	6.2	594	8	09T3A0
31	81.5	6.2	753	5	025050
32	81	6.2	200	6	097912
33	81	6.2	204	2	09LIE4
34	81	6.2	899	2	09R2J7
35	81	6.2	899	2	052336
36	81	6.2	968	10	09LUL8
37	80.5	6.2	173	12	092395
38	80.5	6.2	533	2	09A054
39	80.5	6.2	547	12	0911K0
40	80.5	6.2	549	12	09DX74
41	80.5	6.2	570	10	065813
42	80.5	6.2	888	3	09C209
43	80	6.1	212	6	097698
44	80	6.1	243	11	097895
45	80	6.1	256	6	062670

#### ALIGNMENTS

RESULT 1

066280 ID O66280 PRELIMINARY; PRT; 258 AA.  
AC O66280; 1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE HEAT-LABILE ENTEROTOXIN A SUBUNIT.  
GN LTH A SUBUNIT.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-1032 (ENTEROTOXIGENIC);  
RA Komase K.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-1032 (ENTEROTOXIGENIC);  
RX MEDLINE=9501056; PubMed=7998417;  
RA Tamura S., Asanuma H., Tomita T., Komase K., Kawahara K., Danbara H.,  
Kurata T.;  
RA Hattori N., Watanabe K., Suzuki Y., Nagamine T., Aizawa C., Oya A.,  
a trace amount of the holotoxin as an adjuvant for nasal influenza  
vaccine.";  
RL Vaccine 12:1083-1089(1994).  
DR EMBL; AB011677; BAA25725.1; -.  
DR HSP; P06717; ILTG.  
DR InterPro; IPR001144; Enterotoxin\_A.  
DR InterPro; IPR000886; ER-target.  
DR Pfam; PF01375; Enterotoxin\_A; 1.  
DR PRINTS; PRQ0771; ENTEROTOXIN.  
DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
SQ SEQUENCE 258 AA; 29931 MW; 2BB15D27740EB788 CRC64;

Query Match 98.9%; Score 1290; DB 2; Length 258;  
Best Local Similarity 98.8%; Pred. No. 1.4e-110;  
Matches 237; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NGDRLYRADSPPEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGTFVRYDDGYV 60  
DB 19 NGDKLYRADSPPEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGTFVRYDDGYV 78  
QY 61 STSLSLRAHLAGOSILSGYSTYIYVIATAPNMFNVNDLVGYSPHPYEOEVSALGGIP 120  
DB 79 STSLSLRAHLAGOSILSGYSTYIYVIATAPNMFNVNDLVGYSPHPYEOEVSALGGIP 138  
QY 121 YSQIYGYRVNFGVIDERLHNRREYRDYRNLNIAPAEDGYRLAGPPPDHQAWAREPWI 180  
DB 139 YSQIYGYRVNFGVIDERLHNRREYRDYRNLNIAPAEDGYRLAGPPPDHQAWAREPWI 198  
QY 181 HHAPOGCGNSRTITGTCNEETQNLSTIYLRYQSKVKRQIFSDYQSEVDIYNRIDEL 240  
DB 199 HHAPOGCGNSRTITGTCNEETQNLSTIYLRYQSKVKRQIFSDYQSEVDIYNRIDEL 258

RESULT 2  
Q471182 PRELIMINARY; PRT; 68 AA.  
AC Q47182  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ENTEROTOXIN A (FRAGMENT).  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE OF 64-68 FROM N.A.  
RC STRAIN=H10407;  
RX MEDLINE=83007048; PubMed=6749816;  
RA Yamamoto T., Tamura T., Ryoji M., Kaji A., Yokota T., Takano T.;  
RT "Sequence analysis of the heat-labile enterotoxin subunit B gene  
originating in human enterotoxigenic Escherichia coli.";  
RL J. Bacteriol. 152:506-509(1982).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H10407;  
RX MEDLINE=83114628; PubMed=6759877;  
RA Yamamoto T., Tamura T., Yokota T., Takano T.;  
RT "Overlapping genes in the heat-labile enterotoxin operon originating  
from Escherichia coli human strain.";  
RL Mol. Gen. Genet. 188:356-359(1982).  
RN [3]  
RP -SEQUENCE OF 64-68 FROM N.A.  
RC STRAIN=H10407;  
RX MEDLINE=83265593; PubMed=6348025;  
RA Yamamoto T., Yokota T.;  
RT "Sequence of heat-labile enterotoxin of Escherichia coli pathogenic  
for humans.";  
RL J. Bacteriol. 155:728-733(1983).  
DR EMBL; J01646; AAB02981.1; -;  
DR HSSP; P06717; 1LT3.  
DR InterPro; IPR001144; Enterotoxin\_A.  
DR Pfam; PF01375; Enterotoxin\_A.  
DR PRINTS; PR00771; ENTEROTOXINA.  
DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
FT NON\_TER 1 1  
SQ SEQUENCE 68 AA; 8077 MW; 673CDB4D2E884854 CRC64;

Query Match 27.4%; Score 357; DB 2; Length 68;  
Best Local Similarity 95.6%; Pred. No. 1.2e-25;  
Matches 65; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 173 AWREEPWIHAPOGCGNSRTITGTCNEETQNLSTIYLRYQSKVKRQIFSDYQSEVDI 232

Db 1 AWREEPWIHAPOGCGSSRTITGTCNEETQNLSTIYLRYQSKVKRQIFSDYQSEVDI 60  
QY 233 YNRIDEL 240  
DB 61 YNRIDEL 68  
RESULT 3  
Q57372 PRELIMINARY; PRT; 46 AA.  
AC Q57372;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CHOLERA TOXIN A2.  
GN CTA2.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=42608;  
RX MEDLINE=94237453; PubMed=8181723;  
RA Lebens M., Holmgren J.;  
RT "Structure and arrangement of the cholera toxin genes in Vibrio  
cholerae O139.";  
RL FEMS Microbiol. Lett. 117:197-202(1994).  
DR EMBL; X76391; CAA53975.1; -;  
DR EMBL; X76390; CAA53974.1; -;  
DR HSSP; P01555; 1XTC.  
DR InterPro; IPR001144; Enterotoxin\_A.  
DR InterPro; IPR000886; ER\_target.  
DR Pfam; PF01375; Enterotoxin\_A; 1.  
DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
SQ SEQUENCE 46 AA; 5447 MW; 1B6085A02E8889D6 CRC64;

Query Match 11.7%; Score 152; DB 2; Length 46;  
Best Local Similarity 63.6%; Pred. No. 4.8e-07;  
Matches 28; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 197 DTCNEETQNLSTIYLRYQSKVKRQIFSDYQSEVDIYNRIDEL 240  
DB 3 NTCDERTQSLGKVLDELDEYQSKVKRQYFGYQSDIDTHNRINDEL 46  
RESULT 4  
Q90804 PRELIMINARY; PRT; 850 AA.  
AC Q90804;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PIERISIN.  
OS Pteris forae.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Papilionoidea; Pieridae; Pierinae; Pieris.  
OX NCBI\_TaxID=64459;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99415906; PubMed=10485873;  
RA Watanabe M., Kono T., Matsushima-Hibiya Y., Kanazawa T., Nishisaka N.,  
RA Kishimoto T., Koyama K., Sugimura T., Wakabayashi K.;  
RT "Molecular cloning of an apoptosis-inducing protein, pierisin, from  
cabbage butterfly: Possible involvement of ADP-ribosylation in its  
activity.";  
RT activity.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:10608-10613(1999).  
DR EMBL; AB030305; BAA84491.1; -;  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR Pfam; PF00652; Ricin\_B\_lectin; 8.  
DR SMART; SM00458; RICIN; 4.

DR PROSITE: PS50231; RICIN\_B\_LECTIN; 3.  
SQ SEQUENCE 850 AA; 98082 MW; 2C891698B2C0880E CRC64;

Query Match 7.2%; Score 94; DB 5; Length 850;  
Best Local Similarity 23.5%; Pred. No. 5.7;  
Matches 60; Conservative 32; Mismatches 103; Indels 60; Gaps 11;

QY 4 RLYRADSRPPDEIKRSG--GLMPRGHNEYFDRGTQMNLNLYDHARGTGTGFRVYDDGYVSL 61  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 67 RLVRWDRPPNEIFLDFGVPIVTRNPDWEE-----TDLYGFAKNNHPST-----FVS 114  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 62 TSLSLR-----SAHLAQOSILSGYSTYYIVYVATAPNMFNVNDVLGVSPHPYEQEVS 114  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 115 TTKTORNNKKYVWTPRNRNGIVYOYEIY-----APGGVDVNDSDSPWQMEVA 167  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 115 ALGGIPYSQIYGVYRVNFGVIDERLHRNREYRDRYRNLTAPNMFNVNDVLGVSPHPYEQEVA 174  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 168 FPGQIQNIYRSARELHNGRI-QRIWINPNFLD-----PGDLEPIYSSSTPQVIW 217  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 175 R-BEPWIHHAQ-----GCGNSSRTITGDTCTNEETQNLSTIYLREY--QS 216  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 218 RMNHPDGGHQRDQERSASSYDLMYGGTGN---VQEDTFGDEPNPNKPIAAGEFMIES 273  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 217 KVKRQIFSDYQSEVD 231  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 274 IKDKNSFLDKSNVN 288  
||| ||| ||| ||| : : : : : ||| : : : : :

RESULT 5  
O69258 PRELIMINARY; PRT; 269 AA.  
AC O69258; 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE TOXIN SUBUNIT SI PRECURSOR.  
SL.  
OS Bordetella pertussis.  
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;  
OC Bordetella.  
OX NCBI\_TaxID=520;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=287;  
RA Boursaux-Eude C., Guiso N.;  
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HAV, AND AL1561;  
RA Boursaux-Eude C., Thiberge S., Carletti G., Guiso N.;  
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B572;  
RA Mooli F.R.;  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B572;  
RX MEDLINE=98114370; PubMed=9453625;  
RA Mooli F.R., Vanolschot H., Heuvelman K., vanderHeide H., Gastra W.,  
Willems R.R.J.;  
RT "Polymorphism in the Bordetella pertussis virulence factors  
P.69/pertactin and pertussis toxin in The Netherlands: temporal trends  
and evidence for vaccine-driven evolution.";  
RL Infect. Immun. 66:670-675(1998).  
DR EMBL; AJ006155; CAA06897.1; -.  
DR EMBL; AJ007364; CAA07479.1; -.  
DR EMBL; AJ007363; CAA07478.1; -.  
DR EMBL; AJ245366; CAB51472.1; -.  
DR HSP; P04977; 1BCP.  
DR InterPro; IPR003898; Borpert\_toxa.

DR Pfam: PF02917; Pertussis\_S1; 1.  
KW SIGNAL; Toxin.  
FT SIGNAL 1 34 POTENTIAL.  
FT CHAIN 35 269 TOXIN SUBUNIT S1.  
SQ SEQUENCE 269 AA; 29956 MW; F6CDC9028E6B08AB CRC64;

Query Match 7.1%; Score 93; DB 2; Length 269;  
Best Local Similarity 24.0%; Pred. No. 1.4;  
Matches 53; Conservative 22; Mismatches 62; Indels 84; Gaps 11;

QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNLNLYDHARGTGTGFRVYDDGYVSTSL 64  
:||| ||| ||| ||| : : : : : ||| : : : : :  
Db 41 VYRDSRPPDFQNG-GFTAWGNND-----NVLDLHTGRSCQGVSSNFAVSTSS 89  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 65 SLR-----SAHLAQOSILSGYSTYYIVYVATAPNMFNVNDVLGVSPHPYE 110  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 90 SRRTEVYLEHRMGEAERAGRG--TGFIGIYIEV-RADNNE----- 131  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 111 QEVSALGGIPYSQIYGVYRVNFGVIDERLHRNREYRDRYRNLTAPNMFNVNDVLGVSPHPYE 168  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 132 -----YGAASSYF-----EYVDYGDNAGRILAGA-----LATYQ 161  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 169 PDQAWREEPWIHHAQPCGCGNSSRT-----ITGDTCTNEETQN 205  
:||| ||| ||| ||| : : : : : ||| : : : : :  
Db 162 SEYLAHRRIP-----PENIRRVTRYHNGITGETTTTEYSN 197  
||| ||| ||| ||| : : : : : ||| : : : : :

RESULT 6  
Q9GV36 PRELIMINARY; PRT; 850 AA.  
AC Q9GV36; 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
DE PIERISIN-B.  
OS Pieris brassicae (White butterfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pieris; Pieridae; Pierinae; Pieris.  
OX NCBI\_TaxID=7116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20428462; PubMed=10971585;  
RA Matsushima-Hibiya Y., Watanabe M., Kono T., Kanazawa T., Koyama K.,  
Sugimura T., Wakabayashi K.;  
RT "Purification and cloning of pierisin-2, an apoptosis-inducing protein  
from the cabbage butterfly, pieris brassicae.";  
RL Eur. J. Biochem. 267:5742-5750(2000).  
DR EMBL; AB037676; BAB13774.1; -.  
DR InterPro; IPR000772; Ricin\_B\_lectin.  
DR Pfam; PF00652; Ricin\_B\_lectin; 8.  
DR SMART; SM00458; RICIN; 3.  
DR PROSITE; PS50231; RICIN\_B\_LECTIN; 3.  
SQ SEQUENCE 850 AA; 97988 MW; 6E6044049D36FE73 CRC64;

Query Match 7.1%; Score 93; DB 5; Length 850;  
Best Local Similarity 24.1%; Pred. No. 7.1;  
Matches 62; Conservative 30; Mismatches 101; Indels 64; Gaps 13;

QY 4 RLYRADSRPPDEIKRSG--GLMPRGHNEYFDRGTQMNLNLYDHARGTGTGFRVYDDGYVSL 61  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 67 RLVRWDRPPNEIFLDFGVPIVTRNPDWEE-----TDLYGFAKNNHPST-----FVS 114  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 62 TSLSLR-----SAHLAQOSILSGYSTYYIVYVATAPNMFNVNDVLGVSPHPYEQEVS 114  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 115 TTKTORNNKKYVWTPRNRNGIVYOYEIY-----APGGVDVNDSDSPWQMEVA 167  
||| ||| ||| ||| : : : : : ||| : : : : :  
QY 115 ALGGIPYSQIYGVYRVNFGVIDERLHRNREYRDRYRNLTAPNMFNVNDVLGVSPHPYEQEVA 173  
||| ||| ||| ||| : : : : : ||| : : : : :  
Db 168 FPGQIQ-----NLIYRSARELHNGRI-QRIWINPNFLDGLDLEPIYSSRTIQV 215  
||| ||| ||| ||| : : : : : ||| : : : : :





```
QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRTGMNINLYDHARGTGTGFVYDDGYVSTSL 64
   :|| |||||::: :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 41 VYRDYRPPEDVFQNGFTAGNND-----NVLEHLTGRCQGVSSNSAFVSTSS 89
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 65 SLR-----SAHLAQSLSGYSTYIYVIATAPNMFNVNDVLGVYSPHPYE 110
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 90 SRRYTEVLEHRMQEAVEAERAGRG--TGHFYIGYEV-RADNRF----- 131
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 111 QEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYDRDYRNLN--IAPAECDYRLAGPP 168
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 132 -----YGAASYF-----EYVDYGDNAGRILAGA-----LATYQ 161
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 169 PDQAWREEPWIHAPOGCGNSRST-----ITGDTCTNEETON 205
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 162 SEYLAHRRIP-----PENIRRYRVYHNGITGETTTEYPN 197
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
RESULT 12
Q9NT57 PRELIMINARY; PRT; 547 AA.
ID Q9NT57 AC Q9NT57
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 18, Last annotation update)
DE HYPOTHETICAL 62.2 KDA PROTEIN (FRAGMENT).
GN DKFZP434C0118.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Ootenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI37520; CAB70786.1;
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00646; F-box; 1.
DR PROSITE; PS50181; FBOX; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 547 AA; 62217 MW; A726EE815A82DF31 CRC64;
```

```
Query Match 6.7%; Score 88; DB 4; Length 547;
Best Local Similarity 23.4%; Pred. No. 11;
Matches 39; Conservative 32; Mismatches 62; Indels 34; Gaps 9;

QY 53 VRYDDGVYSTSLRSALHAGQSLSGY-----STYIYVIATAPNMFNVNDVLGVYSP 106
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 364 VREGVVISAKTYIRQ---GEQSLDGFYRAWHQVEYRYI-----RFPDGHVMMLTTP 414
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 107 HPYEQEVSAL---GGIPYSQIYGWYRVN-----FGVIDERL-HRNREYDRDYRNL 153
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 415 EEPQSIYVPLRTRNTRTDALLGHYLSQDNTQTKVFAVITKKKEKPLDYKYRYFRV 474
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 154 NIAPAEQGYR--LAGPPDHQANREEPWIHAPOGCGNSRSTITGDT 198
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 475 PVQEQADQSFHVLQCSGHQRPNKLIWIHNS---CHITYKS-TGET 517
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 13
ID O94527 PRELIMINARY; PRT; 809 AA.
AC O94527
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE GUANINE NUCLEOTIDE BINDING PROTEIN BETA SUBUNIT-LIKE.
GN SPBC609.03.
OS Schizosaccharomyces pombe (Fission yeast).
```

```
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Seeger K., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS):
DR EMBL; AL035226; CAA22832.1; -.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE; PS50082; WD_REPEATS_2; 1.
DR PROSITE; PS50294; WD_REPEATS_REGION; 2.
KW Repeat; WD repeat.
SQ SEQUENCE 809 AA; 92498 MW; 4877443546BD057E CRC64;

Query Match 6.7%; Score 87; DB 3; Length 809;
Best Local Similarity 22.6%; Pred. No. 23;
Matches 43; Conservative 29; Mismatches 54; Indels 64; Gaps 12;

QY 1 NGD---RLYRADSRPPDEIKRSGGLMPRGHNEYFDRTGMNINLYDHARGTGTGFVYDD 57
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 110 SGDKLKLEFLDS-----SKEGGM-----DHGNETQTRCSCALDSVKNIVPCDN 154
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 58 GY-----VSTSLRSALHAGO-----SILSGYSTYI--YVIATAPNMFNVNDVL 101
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 155 GHTFLVCSSEDGTARQYDIREPHVCNODLPCPSILVNYNPNRYNLTYITMSPS----- 206
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 102 GYVSPHPYEQEVSALGGI--PYSQIYGWYRVNFGVIDERLHRNREYDRDYRNLNAPAE 160
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 207 -----NPY---YFAIGGTHPYAFLY-----DRRMVK-KSFEDDM--TWMTSPK 245
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 161 GYRLAGFPDP 170
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 246 CRCVRKFSFD 255
   :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

RESULT 14
Q9JZ10 PRELIMINARY; PRT; 145 AA.
ID Q9JZ10 AC Q9JZ10
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN NMB1343.
GN NMB1343.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; Pubmed=107110307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair H., Cittone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002482; AAF41718.1; -.
DR TIGR; NMB1343; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 145 AA; 16339 MW; 900382FF572BAB3B CRC64;
```

```

Query Match      6.6%; Score 85.5; DB 6; Length 256;
Best Local Similarity 25.6%; Pred. No. 6.6;
Matches 32; Conservative 20; Mismatches 48; Indels 25; Gaps 7;

Y 113 VSALGSGIPSQIYGWYRVNFGVIDERLHNRREYRDYYR-NLNTAPAEDGYLACGFPDH 171
      |||||      : : : : ||||| : : : : |||||
D 125 VGS LGG-----YMLGSAMSRPLIHFGNDYEDYRNNYRPNQVYYR----PYDH 171
      |||||      : : : : ||||| : : : : |||||

Y 172 QAWREEPWTHHAPOCGN---SSRTITGDTCNEQTLSYTREYQSKVRQIFSDYS 228
      |||||      : : : : ||||| : : : : |||||
D 172 YS-NQDNFVH----DCVNITVQKHVTYTTTKGE---NFTETDIKIRVEVQMCITQYOR 223
      |||||      : : : : ||||| : : : : |||||

Y 229 EVDYI 233
      ||
D 224 ESQAY 228

```

Search completed: June 10, 2002, 17:33:06  
Job time: 133 sec





INTERFERENCE

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:15:38 ; Search time 12.93 seconds  
(without alignments)  
453.375 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 1304  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNIRDEL 240

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1299	99.6	240	4	US-08-823-120-3
2	1183	90.7	236	4	US-08-823-120-6
3	1183	90.7	254	3	US-08-486-099-114
4	1183	90.7	254	3	US-08-360-107A-124
5	1183	90.7	254	3	US-08-484-223B-114
6	1183	90.7	254	3	US-08-919-597-114
7	1183	90.7	254	3	US-08-475-668A-114
8	1183	90.7	254	3	US-08-485-551A-114
9	1183	90.7	254	3	US-08-471-913A-114
10	1183	90.7	254	4	US-08-485-264A-114
11	1183	90.7	254	4	US-08-474-349A-114
12	1088	83.4	240	2	US-08-435-605A-13
13	1088	83.4	240	4	US-08-823-120-4
14	1088	83.4	240	4	US-08-823-120-8
15	1088	83.4	258	1	US-08-449-045C-2
16	1088	83.4	258	2	US-08-435-605A-2
17	1088	83.4	258	6	5223610-2
18	1085	83.2	240	2	US-08-435-605A-15
19	1085	83.2	240	2	US-08-435-605A-51
20	1085	83.2	240	2	US-08-435-605A-56
21	1084	83.1	240	2	US-08-435-605A-52
22	1084	83.1	240	2	US-08-435-605A-53
23	1081	82.9	240	2	US-08-435-605A-49
24	1081	82.9	240	2	US-08-435-605A-50
25	972	74.5	238	4	US-08-823-120-2
26	940	72.1	194	2	US-08-435-605A-54
27	934	71.6	194	2	US-08-435-605A-6

28	934	71.6	194	2	US-08-435-605A-14	Sequence 14, Appl
29	933	71.5	192	2	US-08-435-605A-8	Sequence 8, Appl
30	931	71.4	194	2	US-08-435-605A-16	Sequence 16, Appl
31	925	70.9	194	2	US-08-435-605A-55	Sequence 55, Appl
32	871	66.8	179	2	US-08-435-605A-57	Sequence 57, Appl
33	638.5	49.0	241	4	US-08-823-120-1	Sequence 1, Appl
34	157	12.0	46	2	US-08-435-605A-10	Sequence 10, Appl
35	103	7.9	20	1	US-08-171-299B-11	Sequence 11, Appl
36	96	7.4	20	1	US-08-171-299B-10	Sequence 10, Appl
37	93	7.1	266	6	5223610-5	Patent No. 5223610
38	89	6.8	269	6	5433945-5	Patent No. 5433945
39	88	6.7	1584	4	US-09-251-645-6	Sequence 6, Appl
40	86	6.6	15	3	US-08-296-848A-2	Sequence 2, Appl
41	83.5	6.4	687	1	US-08-204-656B-10	Sequence 10, Appl
42	82.5	6.3	417	4	US-08-462-467B-6	Sequence 6, Appl
43	82.5	6.3	417	4	US-08-462-467B-10	Sequence 10, Appl
44	82.5	6.3	556	4	US-08-462-467B-24	Sequence 24, Appl
45	82.5	6.3	556	4	US-08-462-467B-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1  
US-08-823-120-3  
; Sequence 3, Application US/08823120  
; Patent No. 6149919  
; GENERAL INFORMATION:  
; APPLICANT: Domenighini, Mario  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of  
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and  
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,120  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,003  
; FILING DATE: 11-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0315.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-823-120-3.

Query Match 99.6%; Score 1299; DB 4; Length 240;  
Best Local Similarity 99.6%; Pred. No. 1.1e-142;

```
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NGDRLYRADSRPPDEIKRSGLMPRGHNEYFDRGTQNMNINLYDHARGTGTGFVRYDDGYV 60
Db 1 NGDRLYRADSRPPDEIKRSGLMPRGHNEYFDRGTQNMNINLYDHARGTGTGFVRYDDGYV 60
Qy 61 STLSLSRAHLAGOSILSGYSTYYIYVIATAPNMFNNDVLGVYSPHPYEQEVSALGGIP 120
Db 61 STLSLSRAHLAGOSILSGYSTYYIYVIATAPNMFNNDVLGVYSPHPYEQEVSALGGIP 120
Qy 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNLNIAPAEDGYRLAGPPDQHAWEPEWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNLNIAPAEDGYRLAGPPDQHAWEPEWI 180
Qy 181 HPAQCGGNSRTTGTCTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
Db 181 HPAQCGGNSRTTGTCTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
RESULT 2
US-08-823-120-6
; Sequence 6, Application US/08823120
; Patent No. 6149919
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESS: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823,120
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,003
; FILING DATE: 11-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-823-120-6
Query Match 90.7%; Score 1183; DB 4; Length 236;
Best Local Similarity 93.3%; Pred. No. 3.2e-129;
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;
Qy 1 NGDRLYRADSRPPDEIKRSGLMPRGHNEYFDRGTQNMNINLYDHARGTGTGFVRYDDGYV 60
Db 1 NGDRLYRADSRPPDEIKRSGLMPRG-NEYFDRGTQNMNINLYDHARGTGTGFVRYDDGYV 59
```

```
Qy 61 STLSLSRAHLAGOSILSGYSTYYIYVIATAPNMFNNDVLGVYSPHPYEQEVSALGGIP 120
Db 60 STLSLSRAHLAGOYILSGSLTYIVIA---NMFNNDVIVSVSPHPYEQEVSALGGIP 116
Qy 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNLNIAPAEDGYRLAGPPDQHAWEPEWI 180
Db 117 YSQIYGWYRVNFGVIDERLHRNREYRDYRNLNIAPAEDGYRLAGPPDQHAWEPEWI 176
Qy 181 HPAQCGGNSRTTGTCTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
Db 177 HPAQCGGNSRTTGTCTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 236
RESULT 3
US-08-486-099-114
; Sequence 114, Application US/08486099
; Patent No. 6013263
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
; TITLE OF INVENTION: B VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,099
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-031
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-486-099-114
Query Match 90.7%; Score 1183; DB 3; Length 254;
Best Local Similarity 93.3%; Pred. No. 3.6e-129;
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;
Qy 1 NGDRLYRADSRPPDEIKRSGLMPRGHNEYFDRGTQNMNINLYDHARGTGTGFVRYDDGYV 60
Db 19 NGDRLYRADSRPPDEIKRFRSLMPRG-NEYFDRGTQNMNINLYDHARGTGTGFVRYDDGYV 77
Qy 61 STLSLSRAHLAGOSILSGYSTYYIYVIATAPNMFNNDVLGVYSPHPYEQEVSALGGIP 120
```

Db 78 STSLSLSAHLAGQYILSGYSLTIYVIA---NMFNVDIVSYSPHPYQEVSAALGGIP 134  
QY 121 YSQIYGWRYNFGVIDERLHRNREYDRYRNINLAPAEADGYRLAGPPDPHQAWAREPWI 180  
Db 135 YSQIYGWRYNFGVIDERLHRNREYDRYRNINLAPAEADGYRLAGPPDPHQAWAREPWI 194  
QY 181 HHAPOGGCNSRRTTGTCTCNEETQNLSTIYLREYQSVKVKRQIFSDYQSEVDIYNRIDEL 240  
Db 195 HHAPOGGCDSRRTTGTCTCNEETQNLSTIYLREYQSVKVKRQIFSDYQSEVDIYNRIDEL 254

RESULT 4  
US-08-360-107A-124  
; Sequence 124, Application US/08360107A  
; Patent No. 6017536  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION  
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TITLE OF INVENTION: TRANSMISSION  
; NUMBER OF SEQUENCES: 149  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/360,107A  
; FILING DATE: 20-DEC-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-013  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 124:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-360-107A-124

Query Match 90.7%; Score 1183; DB 3; Length 254;  
Best Local Similarity 93.3%; Pred. No. 3.6e-129;  
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;  
QY 1 NGRDLRADSRPDEIKRSGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 60  
Db 19 NGRDLRADSRPDEIKRFRSLMPRG-NEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 77  
QY 61 STSLSLSAHLAGQSLSGSYSTIYVVIATAPNMFNVDIVSYSPHPYQEVSAALGGIP 120  
Db 78 STSLSLSAHLAGQYILSGYSLTIYVIA---NMFNVDIVSYSPHPYQEVSAALGGIP 134

QY 121 YSQIYGWRYNFGVIDERLHRNREYDRYRNINLAPAEADGYRLAGPPDPHQAWAREPWI 180  
Db 135 YSQIYGWRYNFGVIDERLHRNREYDRYRNINLAPAEADGYRLAGPPDPHQAWAREPWI 194  
QY 181 HHAPOGGCNSRRTTGTCTCNEETQNLSTIYLREYQSVKVKRQIFSDYQSEVDIYNRIDEL 240  
Db 195 HHAPOGGCDSRRTTGTCTCNEETQNLSTIYLREYQSVKVKRQIFSDYQSEVDIYNRIDEL 254

RESULT 5  
US-08-484-223B-114  
; Sequence 114, Application US/08484223B  
; Patent No. 6020459  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TITLE OF INVENTION: TRANSMISSION  
; NUMBER OF SEQUENCES: 245  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,223B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-029  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-484-223B-114

Query Match 90.7%; Score 1183; DB 3; Length 254;  
Best Local Similarity 93.3%; Pred. No. 3.6e-129;  
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;  
QY 1 NGRDLRADSRPDEIKRSGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 60  
Db 19 NGRDLRADSRPDEIKRFRSLMPRG-NEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 77  
QY 61 STSLSLSAHLAGQSLSGSYSTIYVVIATAPNMFNVDIVSYSPHPYQEVSAALGGIP 120  
Db 78 STSLSLSAHLAGQYILSGYSLTIYVIA---NMFNVDIVSYSPHPYQEVSAALGGIP 134  
QY 121 YSQIYGWRYNFGVIDERLHRNREYDRYRNINLAPAEADGYRLAGPPDPHQAWAREPWI 180

.....

Db 195 HHAQGGDSSRTITGDTTCNEETQNLSTIYLREYQSKVKROI FSDYOSEVDIYNRIRDEL 254

## RESULT 9

```

US-08-471-913A-114
; Sequence 114, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,913A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-030  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-471-913A-114

Query Match 90.7% Score 1183; DB 3; Length 254;  
 Best Local Similarity 93.3%;  
 Pred. No. 3.6e-129;  
 Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

Qy	1	NGDRLYRADSRPDEIKRSGLMPRGHNEYFDRGTOMNINLNDHARGTQTGFVRDDGVV	60
Db	19	NGDRLYRADSRPDEIKRFXSLPRG-NEYFDRGTOMNINLNDHARGTQTGFVRDDGVV	77
Qy	61	STLSLSRSAHLAQSLGSLGYSTYYIVYIATAPNFMNVNDVLGYISPHPYEQVSVLGGIP	120
Db	78	STLSLSRSAHLAQYILSGYSLYTVIA---NMFMNVNDVSVYSPHPYEQVSVLGGIP	134
Qy	121	YSQIYGYRWYVFGVIDERLHRNEYRDRYYRNUNLINAPEADGYVLGAFPDPDHOAREPWI	180
Db	135	YSQIYGYRWYVFGVIDERLHRNEYRDRYYRNUNLINAPEADGYVLGAFPDPDHOAREPWI	194
Qy	181	HHAPQGCNSRSRTITGDTCNEETQNLSLYLREYQSKVRQIFSDYQSEVDYINRIDEL	240
Db	195	HHAPQGCDSRSRTITGDTCNEETQNLSLYLREYQSKVRQIFSDYQSEVDYINRIDEL	254
RESULT	10		

RESULT 10

Db 195 HHAQGGDSSRTITGDTTCNEETQNLSTIYLREYQSKVKROI FSDYOSEVDIYNRIRDEL 254

RESULT 8  
US-08-485-551A-114  
; Sequence 114, Application US/08485551A  
; Patent No. 6068973  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE  
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS  
; TITLE OF INVENTION: TRANSMISSION

NUMBER OF SEQUENCES: 211  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,551A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-485-551A-114

Query Match 90.7%; Score 1183; DB 3; Length 254;  
Best Local Similarity 93.3%; Pred. No. 3.6e-129;  
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

Qy	1	NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV	60
Db	19	NGDRLYRADSRPPDEIKRFSMLPRG-NEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV	77
Qy	61	STSLSLSAHLAGQSIILSGYSTYYIVYATAPNMFNVNDVLGVSPHPYEQEVSALGGIP	120
Db	78	STSLSLSAHLAGQSIILSGYSTYYIVYATAPNMFNVNDVLGVSPHPYEQEVSALGGIP	134
Qy	121	YSQIYGYRVNFGVIDERLHRNREYRDYRYRNLNIAPAEDGYRLAGFPDPDQHWREEPWI	180
Db	135	YSQIYGYRVNFGVIDERLHRNREYRDYRYRNLNIAPAEDGYRLAGFPDPDQHWREEPWI	194
Qy	181	HHAPQGCNSRSTTGTDCNTEONLSIIYUREYQSKVKRQIFSDYQSEVDIYNRIREL	240
Db	195	HHAPQGCDSRSTTGTDCNTEONLSIIYUREYQSKVKRQIFSDYQSEVDIYNRIREL	254

Db 195 HHAPQCGDSSRTITGDTTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 254

US-08-485-264A-114  
; Sequence 114, Application US/08485264A  
; Patent No. 6228983

GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.

; APPLICANT: Matthews, Thomas J.

; APPLICANT: Wild, Carl T.

; APPLICANT: Barney, Shawn O.

; APPLICANT: Lambert, Dennis M.

; APPLICANT: Petteway, Stephen R.

; APPLICANT: Langlois, Alphonse J.

; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF

; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING

; NUMBER OF SEQUENCES: 232

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/485,264A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7872-021

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 114:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 254 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-485-264A-114

Query Match 90.7%; Score 1183; DB 4; Length 254;

Best Local Similarity 93.3%; Pred. No. 3.6e-129;

Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

QY 1 NGRDLRADSRPPDEIKRSGGLMPRGHNEVFDRTGTMNINLYDHARGTQTGFVRYDDGYV 60

Db 19 NGRDLRADSRPPDEIKRFRSLMPRG-NEVFDRTGTMNINLYDHARGTQTGFVRYDDGYV 77

QY 61 STSLSRSAHLAQSGILSGYSTIYIYVIATAPNMFNVDVLGYSPHPYEQEVSALGGIP 120

Db 78 STSLSRSAHLAQGYILSGYSLTIYIVIA---NMFNVDVSYSPHPYEQEVSALGGIP 134

QY 121 YSOIYGWYRVNFGVIDERLHNRREYRDYRNINIAPAEDGYRLAGFPDPHQAWREPW 180

Db 135 YSOIYGWYRVNFGVIDERLHNRREYRDYRNINIAPAEDGYRLAGFPDPHQAWREPW 194

QY 181 HHAPOCGGSSRTTGTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240

Db 195 HHAPOCGGSSRTTGTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 254

RESULT 11

US-08-474-349A-114

; Sequence 114, Application US/08474349A

; Patent No. 6333395

GENERAL INFORMATION:

; APPLICANT: Bolognesi, Dani P.

; APPLICANT: Matthews, Thomas J.

; APPLICANT: Wild, Carl T.

; APPLICANT: Barney, Shawn O.

; APPLICANT: Lambert, Dennis M.

; APPLICANT: Petteway, Stephen R.

; APPLICANT: Langlois, Alphonse J.

; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE

; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA

; NUMBER OF SEQUENCES: 517

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/474,349A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7872-024

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 114:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 254 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-474-349A-114

Query Match 90.7%; Score 1183; DB 4; Length 254;

Best Local Similarity 93.3%; Pred. No. 3.6e-129;

Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

QY 1 NGRDLRADSRPPDEIKRSGGLMPRGHNEVFDRTGTMNINLYDHARGTQTGFVRYDDGYV 60

Db 19 NGRDLRADSRPPDEIKRFRSLMPRG-NEVFDRTGTMNINLYDHARGTQTGFVRYDDGYV 77

QY 61 STSLSRSAHLAQSGILSGYSTIYIYVIATAPNMFNVDVLGYSPHPYEQEVSALGGIP 120

Db 78 STSLSRSAHLAQGYILSGYSLTIYIVIA---NMFNVDVSYSPHPYEQEVSALGGIP 134

QY 121 YSOIYGWYRVNFGVIDERLHNRREYRDYRNINIAPAEDGYRLAGFPDPHQAWREPW 180

Db 135 YSOIYGWYRVNFGVIDERLHNRREYRDYRNINIAPAEDGYRLAGFPDPHQAWREPW 194

QY 181 HHAPOCGGSSRTTGTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240

Db 195 HHAPOCGGSSRTTGTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 254

RESULT 12

US-08-435-605A-13

; Sequence 13, Application US/08435605A

; Patent No. 5874287

; GENERAL INFORMATION:



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/823,120  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/256,003  
FILING DATE: 11-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: McClung, Barbara G.  
REGISTRATION NUMBER: 33,113  
REFERENCE/DOCKET NUMBER: 0315,001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 601-2708  
TELEFAX: (510) 655-3542  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-823-120-8

Query Match 83.4%; Score 1088; DB 4; Length 240;  
Best Local Similarity 81.7%; Pred. No. 3.5e-118;  
Matches 196; Conservative 24; Mismatches 20; Indels 0; Gaps 0;  
QY 1 NGRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARTQTGTFVRYDDGYV 60  
Db 1 NDDKLYRADSRPPDEIKRSGGLMPRGQSEYFDRGTQMNINLYDHARTQTGTFVRHDDGYV 60  
QY 61 STSLSRSAHLAGOSILSGYSTYIYVIATAPNMFNVNDVLGYSPHPYEQEVSALGGIP 120  
Db 61 STSLSRSAHLVGQITLSGHSTYIYVIATAPNMFNVNDVLGYSPHPDEQVSALGGIP 120  
QY 121 YSQIYGYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGFPDPHQAWREEPWI 180  
Db 121 YSQIYGYRVHFGVLDQLHRNRYRDYRNINLAPAEADGYRLAGFPDPHQAWREEPWI 180  
QY 181 HHAPOCGNSRRTTGTCNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIKDEL 240  
Db 181 HHAPPGCGNAPRSSISNTCDEKTSQSLGVKFLDEYQSKVKRQIFSGYQSDIDTHNRIKDEL 240

RESULT 15  
US-08-449-045C-2  
Sequence 2, Application US/08449045C  
Patent No. 5770203  
GENERAL INFORMATION:  
APPLICANT: Burnette, Neal W.  
APPLICANT: Kaslow, Harvey R.  
TITLE OF INVENTION: RECOMBINANT DNA-DERIVED CHOLERA TOXIN  
TITLE OF INVENTION: SUBUNIT ANALOGS  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Amgen Inc.  
STREET: 1840 De Havilland Drive  
CITY: Thousand Oaks  
STATE: California  
COUNTRY: USA  
ZIP: 91320-1789  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,045C  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
FILING DATE: 02-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/271,222  
FILING DATE: 06-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mazza, Richard J.  
REGISTRATION NUMBER: 27,657  
REFERENCE/DOCKET NUMBER: A-196C  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 258 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-449-045C-2

Query Match 83.4%; Score 1088; DB 1; Length 258;  
Best Local Similarity 81.7%; Pred. No. 3.9e-118;  
Matches 196; Conservative 24; Mismatches 20; Indels 0; Gaps 0;  
QY 1 NGRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARTQTGTFVRYDDGYV 60  
Db 19 NDDKLYRADSRPPDEIKRSGGLMPRGQSEYFDRGTQMNINLYDHARTQTGTFVRHDDGYV 78  
QY 61 STSLSRSAHLAGOSILSGYSTYIYVIATAPNMFNVNDVLGYSPHPYEQEVSALGGIP 120  
Db 79 STSLSRSAHLVGQITLSGHSTYIYVIATAPNMFNVNDVLGYSPHPDEQVSALGGIP 138  
QY 121 YSQIYGYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGFPDPHQAWREEPWI 180  
Db 139 YSQIYGYRVHFGVLDQLHRNRYRDYRNINLAPAEADGYRLAGFPDPHQAWREEPWI 198  
QY 181 HHAPOCGNSRRTTGTCNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIKDEL 240  
Db 199 HHAPPGCGNAPRSSISNTCDEKTSQSLGVKFLDEYQSKVKRQIFSGYQSDIDTHNRIKDEL 258

Search completed: June 10, 2002, 17:32:05  
Job time: 987 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:31:48 ; Search time 108.72 Seconds  
(without alignments)  
776.999 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 1304  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNRDEL 240

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents\_AA.Main:\*

1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pcp.\*

2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pcp.\*

3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pcp.\*

4: /cgn2\_6/ptodata/2/paa/US080\_COMB.pcp.\*

5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pcp.\*

6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pcp.\*

7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pcp.\*

8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pcp.\*

9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pcp.\*

10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pcp.\*

11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pcp.\*

12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pcp.\*

13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pcp.\*

14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pcp.\*

15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pcp.\*

16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pcp.\*

17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pcp.\*

18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pcp.\*

19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pcp.\*

20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pcp.\*

21: /cgn2\_6/ptodata/2/paa/US097\_COMB.pcp.\*

22: /cgn2\_6/ptodata/2/paa/US098\_COMB.pcp.\*

23: /cgn2\_6/ptodata/2/paa/US099\_COMB.pcp.\*

24: /cgn2\_6/ptodata/2/paa/US100\_COMB.pcp.\*

25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pcp.\*

26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1304	100.0	240	16	US-09-297-171-1
2	1299	99.6	240	6	US-08-256-003-3
3	1294	99.2	240	22	US-09-819-917-7
4	1287	98.7	259	1	PCT-US99-30747-4
5	1287	98.7	259	18	US-09-470-124-4
6	1183	90.7	236	6	US-08-256-003-6
7	1183	90.7	236	14	US-09-044-696-2

8	1183	90.7	254	7	US-08-360-107-124	Sequence 124, App
9	1183	90.7	254	8	US-08-470-896-114	Sequence 114, App
10	1183	90.7	254	8	US-08-471-913-114	Sequence 114, App
11	1183	90.7	254	8	US-08-475-668-114	Sequence 114, App
12	1183	90.7	254	8	US-08-484-223-114	Sequence 114, App
13	1183	90.7	254	8	US-08-484-223A-114	Sequence 114, App
14	1183	90.7	254	8	US-08-485-546-114	Sequence 114, App
15	1183	90.7	254	8	US-08-485-546A-114	Sequence 114, App
16	1183	90.7	254	8	US-08-485-551-114	Sequence 114, App
17	1183	90.7	254	8	US-08-487-266-114	Sequence 114, App
18	1183	90.7	254	8	US-08-487-266A-114	Sequence 114, App
19	1183	90.7	254	8	US-08-487-355-114	Sequence 114, App
20	1183	90.7	254	8	US-08-487-355A-114	Sequence 114, App
21	1183	90.7	254	13	US-08-919-600-114	Sequence 114, App
22	1183	90.7	254	19	US-09-502-445-114	Sequence 114, App
23	1088	83.4	240	6	US-08-256-003-4	Sequence 8, Appli
24	1088	83.4	240	6	US-08-256-003-8	Sequence 8, Appli
25	1088	83.4	240	14	US-09-044-696-4	Sequence 4, Appli
26	1088	83.4	240	22	US-09-819-917-8	Sequence 8, Appli
27	1088	83.4	258	1	PCT-US99-30747-46	Sequence 46, Appl
28	1088	83.4	258	18	US-09-470-124-46	Sequence 46, Appl
29	1088	83.4	382	1	PCT-US01-08582-3	Sequence 3, Appli
30	1087	83.4	258	1	PCT-US99-30747-50	Sequence 50, Appl
31	1087	83.4	258	18	US-09-470-124-50	Sequence 48, Appl
32	1084	83.1	258	1	PCT-US99-30747-48	Sequence 48, Appl
33	1084	83.1	382	18	US-09-470-124-48	Sequence 48, Appl
34	1084	83.1	382	1	PCT-US01-08582-4	Sequence 4, Appli
35	1081	82.9	258	1	PCT-US99-30747-52	Sequence 52, Appl
36	1081	82.9	258	18	US-09-470-124-52	Sequence 52, Appl
37	983	75.4	236	22	US-09-819-917-6	Sequence 6, Appli
38	972	74.5	238	6	US-08-256-003-2	Sequence 2, Appli
39	638.5	49.0	241	6	US-08-256-003-1	Sequence 1, Appli
40	638.5	49.0	241	22	US-09-819-917-5	Sequence 5, Appli
41	252.5	19.4	1338	18	US-09-402-100-4	Sequence 4, Appli
42	245	18.8	750	18	US-09-402-100-2	Sequence 2, Appli
43	241.5	18.5	337	21	US-09-760-234-2	Sequence 2, Appli
44	241.5	18.5	461	14	US-09-051-315-2	Sequence 2, Appli
45	241.5	18.5	461	18	US-09-423-493-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-297-171-1

; Sequence 1, Application US/09297171

; GENERAL INFORMATION:

; APPLICANT: Piza, Mariagrazia

; APPLICANT: Giuliani, Marzia M

; APPLICANT: Rappuoli, Rino

; TITLE OF INVENTION: IMMUNOGENIC DETOXIFIED MUTANT E. COLI LT-A TOXIN

; FILE REFERENCE: 0342.001 / 2302-0342

; CURRENT APPLICATION NUMBER: US/09/297,171

; CURRENT FILING DATE: 1999-04-27

; PRIOR APPLICATION NUMBER: PCT/IB97/01440

; PRIOR FILING DATE: 1997-10-30

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 1

; LENGTH: 240

; TYPE: PRT

; ORGANISM: Escherichia coli

US-09-297-171-1

Query Match 100.0%; Score 1304; DB 16; Length 240;

Best Local Similarity 100.0%; Pred. No. 5 5e-137;

Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60

Db 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60

```
QY 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
Db 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
QY 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
QY 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
Db 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240

RESULT 2
US-08-256-003-3
; Sequence 3, Application US/08256003
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,003
; FILING DATE: 11-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315,001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-256-003-3

Query Match 99.6%; Score 1299; DB 6; Length 240;
Best Local Similarity 99.6%; Pred. No. 2e-136;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGDRLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60
Db 1 NGDRLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60
QY 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
Db 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
QY 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
QY 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
Db 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240

RESULT 3
US-09-819-917-7
; Sequence 7, Application US/09819917
; GENERAL INFORMATION:
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Fontana, Maria Rita
; APPLICANT: Giannelli, Valentina
; APPLICANT: Rappuoli, Rina
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of Cholera Toxin
; FILE REFERENCE: CHIR0312
; CURRENT APPLICATION NUMBER: US/09/819,917
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 08/981,208
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 9513371.6
; PRIOR FILING DATE: 1995-06-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 240
; TYPE: PRT
; ORGANISM: E. coli
US-09-819-917-7

Query Match 99.2%; Score 1294; DB 22; Length 240;
Best Local Similarity 99.2%; Pred. No. 7.3e-136;
Matches 238; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGDRLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60
Db 1 NGDRLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60
QY 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
Db 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
QY 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
QY 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
Db 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240

RESULT 4
PCT-US99-30747-4
; Sequence 4, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4869/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
PCT-US99-30747-4
```

```
QY 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
Db 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
QY 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
QY 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
Db 181 HHAPQCGGNSRRTITGDTCTNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240

RESULT 2
US-08-256-003-3
; Sequence 3, Application US/08256003
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,003
; FILING DATE: 11-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315,001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-256-003-3

Query Match 99.6%; Score 1299; DB 6; Length 240;
Best Local Similarity 99.6%; Pred. No. 2e-136;
Matches 239; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGDRLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60
Db 1 NGDRLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTGFVRYDDGYV 60
QY 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
Db 61 STSLRSALHLAGOSILSGYSTYYIVYIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP 120
QY 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNINLAPAEADGYRLAGPPDHOAWREEPWI 180
```

```
Query Match      98.7%; Score 1287; DB 1; Length 259;
Best Local Similarity 98.3%; Pred. No. 5e-135;
Matches 236; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 60
DB 20 NGDKLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 79

QY 61 STSLSRSAHLAQSILSGYSTYIIYVIATAPNMFNVDVLYVSPHPYEQEVSALGGIP 120
DB 80 STSLSRSAHLAQSILSGYSTYIIYVIATAPNMFNVDVLYVSPHPYEQEVSALGGIP 139

QY 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 180
DB 140 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 199

QY 181 HHAPOCGGSSRITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
DB 200 HHAPOCGGSSRITGDTCTNEETQNLSTIYLYRKYQSKVKRQIFSDYQSEVDIYNRIRDEL 259

RESULT 5
US-09-470-124-4
; Sequence 4, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; APPLICANT: Arintzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
US-09-470-124-4

Query Match      98.7%; Score 1287; DB 18; Length 259;
Best Local Similarity 98.3%; Pred. No. 5e-135;
Matches 236; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 60
DB 20 NGDKLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 79

QY 61 STSLSRSAHLAQSILSGYSTYIIYVIATAPNMFNVDVLYVSPHPYEQEVSALGGIP 120
DB 80 STSLSRSAHLAQSILSGYSTYIIYVIATAPNMFNVDVLYVSPHPYEQEVSALGGIP 139

QY 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 180
DB 140 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 199

QY 181 HHAPOCGGSSRITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
DB 200 HHAPOCGGSSRITGDTCTNEETQNLSTIYLYRKYQSKVKRQIFSDYQSEVDIYNRIRDEL 259

RESULT 6
US-08-256-003-6
; Sequence 6, Application US/08256003
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
```

```
; APPLICANT: Pizza, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94508-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,003
; FILING DATE: 11-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-256-003-6

Query Match      90.7%; Score 1183; DB 6; Length 236;
Best Local Similarity 93.3%; Pred. No. 2e-123;
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;

QY 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 60
DB 1 NGDRLYRADSRPPDEIKRPSLMPRG-NEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 59

QY 61 STSLSRSAHLAQSILSGYSTYIIYVIATAPNMFNVDVLYVSPHPYEQEVSALGGIP 120
DB 60 STSLSRSAHLAQSILSGYSTIYVIA---NMFNVDVLSVSPHPYEQEVSALGGIP 116

QY 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 180
DB 117 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 176

QY 181 HHAPOCGGSSRITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 240
DB 177 HHAPOCGGSSRITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIRDEL 236

RESULT 7
US-09-044-696-2
; Sequence 2, Application US/09044696
; GENERAL INFORMATION:
; APPLICANT: BARCHFELD, GAIL
; APPLICANT: DEL GIUDICE, GIUSEPPE
; APPLICANT: RAPPUOLI, RINO
; TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
; TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
; ADDRESS: R440
; STREET: P.O. BOX 8097
; CITY: EMERYVILLE
```

; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/044, 696  
; FILING DATE: 18-MAR-1993  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 50/041,227  
; FILING DATE: 21-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: HARBIN, ALISA A.  
; REGISTRATION NUMBER: 33,895  
; REFERENCE/DOCKET NUMBER: 1393.002  
; TELEPHONE: (510) 655-8730  
; TELEPHONE: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-044-696-2

Query Match 90.7%; Score 1183; DB 14; Length 236;  
Best Local Similarity 93.3%; Pred. No. 2e-123;  
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;  
QY 1 NGDRLYRADSPPDEIKRSGGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 60  
Db 1 NGDRLYRADSPPDEIKRSGGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 59  
QY 61 STLSLSRAHLAQGYILSGYSLTYIVIA--NMFNVDVSVSPHPYEQEVSALGGIP 120  
Db 60 STLSLSRAHLAQGYILSGYSLTYIVIA--NMFNVDVSVSPHPYEQEVSALGGIP 116  
QY 121 YSQIYGWYRNFVGVDERLHRNREYDRYRNLIAPAEDGYRLAGPPDQAWREPWI 180  
Db 117 YSQIYGWYRNFVGVDERLHRNREYDRYRNLIAPAEDGYRLAGPPDQAWREPWI 176  
QY 181 HHAPOCGNSRRTITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIDEL 240  
Db 177 HHAPOCGNSRRTITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIDEL 236

RESULT 8  
US-08-360-107-124  
; Sequence 124, Application US/08360107  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION  
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; NUMBER OF SEQUENCES: 137  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA

; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/360,107  
; FILING DATE: 20-DEC-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-013  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8664  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 124:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-360-107-124

Query Match 90.7%; Score 1183; DB 7; Length 254;  
Best Local Similarity 93.3%; Pred. No. 2.2e-123;  
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;  
QY 1 NGDRLYRADSPPDEIKRSGGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 60  
Db 19 NGDRLYRADSPPDEIKRSGGLMPRGHNEYFDRGTQNNINLYDHARTQTGFVRYDDGYV 77  
QY 61 STLSLSRAHLAQGYILSGYSLTYIVIA--NMFNVDVSVSPHPYEQEVSALGGIP 120  
Db 78 STLSLSRAHLAQGYILSGYSLTYIVIA--NMFNVDVSVSPHPYEQEVSALGGIP 134  
QY 121 YSQIYGWYRNFVGVDERLHRNREYDRYRNLIAPAEDGYRLAGPPDQAWREPWI 180  
Db 135 YSQIYGWYRNFVGVDERLHRNREYDRYRNLIAPAEDGYRLAGPPDQAWREPWI 194  
QY 181 HHAPOCGNSRRTITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIDEL 240  
Db 195 HHAPOCGNSRRTITGDTCTNEETQNLSTIYLYREYQSKVKRQIFSDYQSEVDIYNRIDEL 254

RESULT 9  
US-08-470-896-114  
; Sequence 114, Application US/08470896  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION  
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; NUMBER OF SEQUENCES: 273  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

Query Match 90.7%; Score 1183; DB 8; Length 254;  
Best Local Similarity 93.3%; Pred. No. 2.2e-123;  
Matches 224; Conservative 2; Mismatches 10; Indels

	Matches	224;	Conservative	2;	Mismatches	10;	Indels	4;	Gaps
Qy	1	NGDRLYRADSRPDEIKRSGGLMPRGHNEYFDRTGTONINLYDHARGTQTGFVRYDDGV	60						
Db	19	NGDRLYRADSRPDEIKRFSRLPRG-NEYFDRTGTONINLYDHARGTQTGFVRYDDGV	77						
Qy	61	STSLSRSAHLAQSIILSGYSTYYIVYIATAPNNFNVDVLGYSPHPYEQEVSALGGIP	120						
Db	78	STSLSRSAHLAQYLILSGYSLIYIVIA---NMFNVNDISVSPHPYEQEVSALGGIP	134						
Qy	121	YSQIGYRVNFVGVIDERLHRNREYDRDYRNLNIAPAEDGYRLAGFPDPHQAWREEPW	180						
Db	135	YSQIGYRVNFVGVIDERLHRNREYDRDYRNLNIAPAEDGYRLAGFPDPHQAWREEPW	194						
Qy	181	HHAPOGCNSRRITGTCTNEETONTSTIYLREYQSXVKRKQIFSDYOSEVDIYNRIDEL	240						
Db	195	HHAPOGCDSRRITGTCTNEETONTSTIYLREYQSXVKRKQIFSDYOSEVDIYNRIDEL	254						

```

RESULT 11
US-08-475-668-114
; Sequence 114, Application US/08475668
; GENERAL INFORMATION:
;
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
;
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
;
; TITLE OF INVENTION: COMPOSITIONS
; TITLE OF INVENTION: MEMBRANE FUSION
; TITLE OF INVENTION: VIRUS TRANSMISSION
;
; NUMBER OF SEQUENCES: 209
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Pennile & Edmonds
; STREET: 1155 Avenue of the Amer
; CITY: New York
; STATE: New York
; COUNTRY: USA
;
; ZIP: 10036-2711
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,668
; FILING DATE: 07-JUN-1995

```

; AFFILIATION: LANGUAGES, RESPONSE U.  
 ; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
 ; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA

CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-026  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-475-668-114

Query Match 90.7%; Score 1183; DB 8; Length 254;  
Best Local Similarity 93.3%; Pred. No. 2.2e-123;  
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;  
QY 1 NGRDLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTGTGFVRYDDGYV 60  
Db 19 NGRDLRADSRPDEIKRFRSLMPRG-NEYFDRGTQMNINLYDHARGTGTGFVRYDDGYV 77  
QY 61 STSLRSALHAGQSLSGSYTYIYVIATAPNMFNVNDVLGYSPHPYEQEVSALGGIP 120  
Db 78 STSLRSALHAGQSLSGSYTYIYVIA---NMFNVNDVISYSPHPYEQEVSALGGIP 134  
QY 121 YSQIYGWYRNFVGVDERLHRNREYDRYRNINIAPAEDGYRLAGFPDPHQAWREPWI 180  
Db 135 YSQIYGWYRNFVGVDERLHRNREYDRYRNINIAPAEDGYRLAGFPDPHQAWREPWI 194  
QY 181 HHAPOCGGSSRTITGTCNEETQNLSTIYLYREYQSKVRQIFSDYQSEVDIYNRIDEL 240  
Db 195 HHAPOCGGSSRTITGTCNEETQNLSTIYLYREYQSKVRQIFSDYQSEVDIYNRIDEL 254

RESULT 12  
US-08-484-223-114  
Sequence 114, Application US/08484223  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 209  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,223  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-029  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-223-114

Query Match 90.7%; Score 1183; DB 8; Length 254;  
Best Local Similarity 93.3%; Pred. No. 2.2e-123;  
Matches 224; Conservative 2; Mismatches 10; Indels 4; Gaps 2;  
QY 1 NGRDLRADSRPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTGTGFVRYDDGYV 60  
Db 19 NGRDLRADSRPDEIKRFRSLMPRG-NEYFDRGTQMNINLYDHARGTGTGFVRYDDGYV 77  
QY 61 STSLRSALHAGQSLSGSYTYIYVIATAPNMFNVNDVLGYSPHPYEQEVSALGGIP 120  
Db 78 STSLRSALHAGQSLSGSYTYIYVIA---NMFNVNDVISYSPHPYEQEVSALGGIP 134  
QY 121 YSQIYGWYRNFVGVDERLHRNREYDRYRNINIAPAEDGYRLAGFPDPHQAWREPWI 180  
Db 135 YSQIYGWYRNFVGVDERLHRNREYDRYRNINIAPAEDGYRLAGFPDPHQAWREPWI 194  
QY 181 HHAPOCGGSSRTITGTCNEETQNLSTIYLYREYQSKVRQIFSDYQSEVDIYNRIDEL 240  
Db 195 HHAPOCGGSSRTITGTCNEETQNLSTIYLYREYQSKVRQIFSDYQSEVDIYNRIDEL 254

RESULT 13  
US-08-484-223A-114  
Sequence 114, Application US/08484223A  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 239  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,223A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-029  
TELECOMMUNICATION INFORMATION:





; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-485-546A-114

Query Match		90.7%	Score 1183;	DB 8;	Length 254;
Best Local Similarity		93.3%	Pred. No. 2.2e-123;		
Matches 224;		Conservative 2;	Mismatches 10;	Indels 4;	Gaps 2;
Qy	1	NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTGTGFVRYDDGYV	60		
Db	19	NGDRLYRADSRPPDEIKRFRSLMPRG-NEYFDRGTOMNINLYDHARGTGTGFVRYDDGYV	77		
Qy	61	STSLSLSAHLAQOSILSGYSTYYIYVIATAPNMFNVNDVLGVYSPHPYEQEVSALGGIP	120		
Db	78	STSLSLSAHLAQGYILSGYSLTIYIVIA---NMFNVDVISYSPHPYEQEVSALGGIP	134		
Qy	121	YSQIYGWYRVNFGVIDERLHRNREYRDYRNLNIAPAEDGYRLAGFPDPHQAWREEPWI	180		
Db	135	YSQIYGWYRVNFGVIDERLHRNREYRDYRNLNIAPAEDGYRLAGFPDPHQAWREEPWI	194		
Qy	181	HHAPQCGGSSRTTGTCTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIDEL	240		
Db	195	HHAPQCGGSSRTTGTCTCNEETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIDEL	254		

Search completed: June 10, 2002, 17:35:18  
Job time: 210 sec





## RESULT

ORGANISM: *Aeropyrum pernix*

## RESULT

```

: IIPB: PRI
:
: ORGANISM: SPHINGOMONAS
:
: FEATURE:
:
: NAME/KEY: unsure
:
: LOCATION: (1)..(875)
:
: OTHER INFORMATION: unsure at all Xaa locations
US-60-360-039-17808

```





; APPLICANT: Williamson, R. Anthony  
; APPLICANT: Burton, Dennis R.  
; TITLE OF INVENTION: ANTIBODIES ABOLISH PRION PROPAGATION AND  
; TITLE OF INVENTION: PROMOTE CLEARANCE OF INFECTIVITY  
; FILE REFERENCE: UCAL-244WO  
; CURRENT APPLICATION NUMBER: PCT/US02/13346  
; CURRENT FILING DATE: 2002-04-25  
; PRIOR APPLICATION NUMBER: 60/287,971  
; PRIOR FILING DATE: 2001-05-01  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: bovine  
PCT-US02-13346-3

Query Match 6.2%; Score 81; DB 1; Length 263;  
Best Local Similarity 24.2%; Pred. No. 3.9;  
Matches 30; Conservative 21; Mismatches 51; Indels 22; Gaps 6;  
QY 114 SALGGIPYSQIYGVYRNFVGVIDERLHRNREYRDYR- NLNAPAEADGYRLAGFPDPHQ 172  
Db 130 AVVGG- - - - -GGYMLGSAMSRLIHFGSDYEDRYRENHRYPNQVYR- - - - -PVDQ 178  
QY 173 AWREEPWIHAPOGCGN- - -SSRTIGDTCNEETQNLSTIYLREYQSKVKRQIFSDYQSE 229  
Db 179 YSNQNNFVH- - - -DCVNITVKEHTVTTTKGE- - -NFTETDIKMMERVVEQMCVTQYQKE 231  
QY 230 VDIY 233  
Db 232 SQAY 235

RESULT 15  
US-09-602-775C-26  
; Sequence 26; Application US/09602775C  
; GENERAL INFORMATION:  
; APPLICANT: Cashman, Neil  
; APPLICANT: Paramithiotis, Eustache  
; APPLICANT: Slon-Usakiewicz, Jacek  
; APPLICANT: Haghighat, Ashkan  
; APPLICANT: Pinard, Marc  
; APPLICANT: Lawton, Trebor  
; TITLE OF INVENTION: PRION PROTEIN PEPTIDES AND USES THEREOF  
; FILE REFERENCE: 50111/002002  
; CURRENT APPLICATION NUMBER: US/09/602,775C  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/140,634  
; PRIOR FILING DATE: 1999-06-23  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 26  
; LENGTH: 264  
; TYPE: PRT  
; ORGANISM: Bos taurus  
US-09-602-775C-26

Query Match 6.2%; Score 81; DB 5; Length 264;  
Best Local Similarity 24.2%; Pred. No. 3.9;  
Matches 30; Conservative 21; Mismatches 51; Indels 22; Gaps 6;  
QY 114 SALGGIPYSQIYGVYRNFVGVIDERLHRNREYRDYR- NLNAPAEADGYRLAGFPDPHQ 172  
Db 131 AVVGG- - - - -GGYMLGSAMSRLIHFGSDYEDRYRENHRYPNQVYR- - - - -PVDQ 179  
QY 173 AWREEPWIHAPOGCGN- - -SSRTIGDTCNEETQNLSTIYLREYQSKVKRQIFSDYQSE 229  
Db 180 YSNQNNFVH- - - -DCVNITVKEHTVTTTKGE- - -NFTETDIKMMERVVEQMCITQYQRE 232  
QY 230 VDIY 233

Db 233 SQAY 236

Search completed: June 10, 2002, 17:35:40  
Job time: 212 sec



---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:32:33 ; Search time 29.32 Seconds  
(without alignments)  
909.199 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 240  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNIRDEL 240

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0  
Searched: 747574 seqs, 111073796 residues

Word size : 8  
Total number of hits satisfying chosen parameters: 82

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :	A_Geneseq_032802.*
1:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	240	100.0	258	6 AAP50190	Sequence encoded b
2	184	76.7	259	21 AAY96646	Plant-optimized E.
3	184	76.7	259	21 AAY96648	Plant-optimized E.
4	179	74.6	258	6 AAP50191	Sequence encoded b
5	177	73.8	382	22 AAU00506	E. coli heat-lab
6	128	53.3	380	22 AAU00507	E. coli heat-lab
7	125	52.1	259	21 AAY96647	Synthetic E. coli
8	125	52.1	259	21 AAY96650	Plant-optimized E.
9	116	48.3	259	21 AAY96649	Plant-optimized E.
10	116	48.3	259	21 AAY96651	Plant-optimized E.
11	111	46.2	240	19 AAW65074	E. coli LT-A prote

12	99	41.2	240	19	AAW65075	E. coli LT-A mutan
13	86	35.8	236	14	AAR44016	"Lys-63" E.coli he
14	86	35.8	236	14	AAR44017	"Lys-97" E.coli he
15	86	35.8	236	14	AAR44018	"Tyr-97" E.coli he
16	86	35.8	236	14	AAR38728	E.coli heat labile
17	86	35.8	236	14	AAR38730	"Asp-53" E.coli he
18	86	35.8	236	14	AAR38731	"Tyr-53" E.coli he
19	86	35.8	236	14	AAR38732	E. coli heat labil
20	86	35.8	237	20	AAW67772	Peptide sequence f
21	86	35.8	254	22	AAU14105	"Lys-104" E.coli h
22	84	35.0	236	14	AAR44020	"Lys-104" E.coli h
23	84	35.0	236	14	AAR44021	"Ser-104" E.coli h
24	84	35.0	236	14	AAR44022	"Ser-104" E.coli h
25	82	34.2	236	14	AAR44023	"Ser-106" E.coli h
26	81	33.8	236	14	AAR44019	"Glu-107" E.coli h
27	74	30.8	236	14	AAR44024	"Glu-114" E.coli h
28	74	30.8	236	14	AAR44025	"Lys-114" E.coli h
29	26	10.8	240	14	AAR44030	Asn-107 cholera to
30	26	10.8	240	14	AAR44027	Lys-63 cholera tox
31	26	10.8	240	14	AAR44028	Lys-97 cholera tox
32	26	10.8	240	14	AAR44029	Ser-106 cholera to
33	26	10.8	240	14	AAR44031	Ser-110 cholera to
34	26	10.8	240	14	AAR44032	Ala-112 cholera to
35	26	10.8	240	14	AAR44033	Glu-114 cholera to
36	26	10.8	240	14	AAR38729	Cholera toxin subu
37	26	10.8	240	19	AAW80809	Amino acid sequenc
38	26	10.8	240	19	AAW80807	Cholera toxin subu
39	26	10.8	240	20	AAW67773	Plant-optimized V.
40	26	10.8	258	21	AAY96653	Plant-optimized mu
41	26	10.8	258	21	AAY96654	Plant-optimized mu
42	26	10.8	258	21	AAY96655	Plant-optimized mu
43	26	10.8	258	21	AAY96656	Plant-optimized mu
44	26	10.8	258	22	AAG65991	Cholera toxin A su
45	25	10.4	258	12	AAR13117	Cholera toxin Al f

ALIGNMENTS

RESULT 1	
AAP50190	
ID AAP50190 standard; Protein; 258 AA.	
XX	
AC AAP50190;	
XX	
DT 30-OCT-1991 (first entry)	
XX	
DE Sequence encoded by the pig scours heat labile toxin (LT) LTA gene.	
XX	
KW Pig scours vaccine; toxin; diarrhoea.	
XX	
OS E.coli NCIB 11932.	
XX	
PN EP145486-A.	
XX	
PD 19-JUN-1985.	
XX	
PF 12-DEC-1984; 84EP-0308620.	
XX	
PR 12-DEC-1983; 83GB-0033131.	
XX	
PA (GLAX ) GLAXO GROUP LTD.	
XX	
PI Hayes MV, Harford S, Ross GW;	
XX	
DR WPI; 1985-148358/25.	
DR N-PSDB; AAN50205.	
XX	
PT New toxoid as inactivated form of toxin for use in vaccines - is	
PT obtd. from organism transformed by gene	
XX	
PS Disclosure; Fig 1; 61pp; English.	
XX	



PF 22-DEC-1999; 99WO-US30747.  
XX  
PR 22-DEC-1998; 98US-0113507.  
XX  
PA (BOYC-) BQYCE THOMPSON INST PLANT RES.  
PA (MASO/) MASON H S.  
PA (ARNT/) ARNTZEN C J.  
XX  
PI Mason HS, Arntzen CJ;  
XX WPI; 2000-442653/38.  
DR N-PSDB; AAA51544.  
XX  
PT New polynucleotides encoding LT-A or CT-A polypeptides for the  
PT transformation of plant cells, useful in immunogenic compositions to  
PT elicit immune responses in animals  
XX  
PS Example 3; Page -: 103pp; English.  
XX  
CC This is mutant R192G Escherichia coli heat-labile toxin (LT) A subunit  
CC (LT-A). The wild-type arginine was replaced with glycine at residue 192  
CC of the mature protein, which was caused by a codon change of TCC to AAG  
CC in the coding sequence. The sequence contains plant-preferred codons and  
CC eliminates sequence motifs associated with spurious mRNA processing. A  
CC single codon insertion (GTG encoding valine) was made to accommodate the  
CC creation of a NcoI restriction site around the initiator methionine  
CC codon. Novel polynucleotides encode a mutant LT-A polypeptide or a  
CC mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A) polypeptide,  
CC which have reduced enzyme activity as compared to the wild-type LT-A or  
CC CT-A polypeptide and where at least one of the codons is altered to a  
CC plant preferred codon. The polynucleotide further comprises a nucleic  
CC acid sequence encoding LT B subunit (LT-B) or a CT B subunit (CT-B). The  
CC polynucleotides are useful for the transformation of plant cells for the  
CC production of transgenic plants to produce edible vaccines, especially  
CC oral vaccines in transgenic plants for the prophylactic or therapeutic  
CC treatment against E. coli or V. cholerae. The mutant polypeptides are  
CC also useful as adjuvants.  
CC NB: This sequence does not appear in the specification, it was made from  
CC the wild type sequence shown in AAY96647, which appears in Figure 1.  
XX  
SQ Sequence 259 AA;

Query Match 76.7%; Score 184; DB 21; Length 259;  
Best Local Similarity 100.0%; Pred. No. 5.4e-181; Indels 0; Gaps 0;  
Matches 184; Conservative 0; Mismatches 0;  
Qy 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRTQMNLINLYDHARGTGTGFVRYDDGVYSTSL 64  
Db 24 lyradsrppdeikrsgglmpRGHNEYfdrgtqminlydhargtqgtfvrYddgyvstsl 83  
Qy 65 SLRSAHLAQQSILSGYSTYIYVIATAPNMFNVNDVLGVSPHPYEQEVSALGGIPYSQI 124  
Db 84 slrsahlagqsilsgystyiyviatapnmfnvndvlgvspHPYeqevsalslgipysq1 143  
Qy 125 YGWRVNVFGVIDERLHRNREYDRYRNINLNTAPADGVRLAGFPPDQHWREEPWIHAP 184  
Db 144 ygwrvnvfgviderlhrnreydryrnrlnlnlntapadgyrlagfppdhqawreepwihap 203  
Qy 185 QGCG 188  
Db 204 qgcg 207  
RESULT 4  
AAP50191  
ID AAP50191 standard; Protein; 258 AA.  
XX  
AC AAP50191;  
XX  
DT 30-OCT-1991 (first entry)  
XX  
DE Sequence encoded by the pig scours heat labile toxin (LT) LTA gene

DE of the site directed mutant SDM1.  
XX Pig scours vaccine; toxin; diarrhoea.  
XX  
OS E.coli NCIB 11932.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 79 /note= "Ser in native SQ"  
XX  
XX EPI45486-A.  
XX  
XX 19-JUN-1985.  
XX  
XX 12-DEC-1984; 84EP-0308620.  
XX  
XX 12-DEC-1983; 83GB-0033131.  
XX  
XX (GLAX ) GLAXO GROUP LTD.  
XX  
XX Hayes MV, Harford S, Ross GW;  
PI WPI; 1985-148358/25.  
DR N-PSDB; AAN50206.  
XX  
XX New toxoid as inactivated form of toxin for use in vaccines - is  
PT obtd. from organism transformed by gene  
XX  
XX Example; Fig 2; 61pp; English.  
XX  
CC AAN50205 is the gene sequence of the natural LTA gene. The LTA gene of  
CC the site directed mutant SDM1 (see AAN50206) is inactive. The  
CC inventors claim a vaccine prepn. active against pig scours which  
CC contains an inactivated LTA component, together with additional K88  
CC antigens opt. with whole cells comprising the antigens or contg. the  
CC inactivated LTA.  
XX  
SQ Sequence 258 AA;  
Query Match 74.6%; Score 179; DB 6; Length 258;  
Best Local Similarity 100.0%; Pred. No. 7.6e-176; Indels 0; Gaps 0;  
Matches 179; Conservative 0; Mismatches 0;  
Qy 62 TSLSRSHLAGQSILSGYSTYIYVIATAPNMFNVNDVLGVSPHPYEQEVSALGGIPY 121  
Db 80 tslsrshlagqsilsgystyiyviatapnmfnvndvlgvspHPYeqevsalslgipY 139  
Qy 122 SQIYGWRVNVFGVIDERLHRNREYDRYRNINLNTAPADGVRLAGFPPDQHWREEPWIH 181  
Db 140 sqlygwrvnvfgviderlhrnreydryrnrlnlnlntapadgyrlagfppdhqawreepwih 199  
Qy 182 HAPQCGNSSRTITGDTCTNEETQNLSTLYLREYQSKVRQIFSDYQSEVDIYNRIRDEL 240  
Db 200 hapqcgnsrrtltgdtcneetqnlstlylreyqskvrqifsdYqsevdynrirdel 258  
RESULT 5  
AAU00506  
ID AAU00506 standard; Protein; 382 AA.  
XX  
AC AAU00506;  
XX  
XX 29-AUG-2001 (first entry)  
XX  
XX E. coli heat-labile enterotoxin (LT) mutant LTS63Y.  
XX  
KW Heat-labile enterotoxin; LT; LTS63Y; Ltgell10/112; mutant;  
KW detoxified and immunologically active protein; ADP-ribosylation; Gs;  
KW endotoxin; diarrhoea; mutein.  
XX  
OS Escherichia coli strain K88ac.  
OS Synthetic.

```
XX FH Key Location/Qualifiers
FT Misc_feature 7
FT /note= "Important residue for enzymatic activity"
FT Misc_feature 44
FT /note= "Important residue for enzymatic activity"
FT Region 58..72
FT /note= "Forms the NAD-binding site"
FT Misc_feature 61
FT /note= "Important residue for enzymatic activity"
FT Misc-difference 63
FT /note= "Substitution of wild type Ser to Tyr"
FT Misc_feature 110
FT /note= "Important residue for enzymatic activity"
FT Misc_feature 112
FT /note= "Important residue for enzymatic activity"
FT Misc-difference 259
FT /note= "Encoded by TG"
XX
XX WO200119998-Al.
XX
XX 22-MAR-2001.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
XX
XX Park EJ, Kim JS, Chang J, Yum J, Chung S;
XX
XX WPI; 2001-281524/29.
XX
XX N-PSDB; AAS01506.
XX
XX New detoxified mutants of Escherichia coli heat-labile enterotoxin
XX useful as vaccine for preventing and treating diarrhoea, and as adjuvant
XX for antibody production
XX
XX Claim 2; Page 39-41; 48pp; English.
XX
XX The present sequence represents Escherichia coli heat-labile
XX enterotoxin (LT) mutant LTS63Y and Ltdell10/112 (AAU00507)
XX are two novel detoxified and immunologically active proteins (LT
XX mutants) derived by site-directed mutagenesis of the A1 subunit of wild
XX type LT. The substitution of Ser to Tyr at position 63 in LTS63Y blocks
XX NAD-binding. Deletion of Glu residues at positions 110 and 112 in
XX Ltdell10/112 eliminate the enzymatic activity of LT. The A1 subunit of
XX wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that
XX regulates cAMP levels. The resulting increase in cAMP is the cause of
XX diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities
XX of mutant heat-labile endotoxins LTS63Y and Ltdell10/112 were tested.
XX Groups of mice were immunised with LTS623Y or Ltdell10/112. The control
XX faecal antibody titres to LT were determined. The results showed that
XX mice immunised with LTS63Y or Ltdell10/112 contained high and
XX comparable level of anti-LT antibodies in sera and faecal extracts
XX compared with those immunised with wild-type LT. The LT mutants are
XX useful as a vaccine for preventing and treating diarrhoea and as an
XX adjuvant for antibody production.
XX
XX Sequence 382 AA;

Query Match 73.8%; Score 177; DB 22; Length 382;
Best Local Similarity 100.0%; Pred. No. 1.2e-173;
Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 LLSRAHLAQGSILSGYSTYIYVIATAPNMFNVNDVLGYSPHPYQEVSAIAGGIPYSQ 123
DB 82 LLSRAHLAQGSILSGYSTYIYVIATAPNMFNVNDVLGYSPHPYQEVSAIAGGIPYSQ 141
QY 124 IYGYRVNFGVIDERLHRNREYDRYYRNINIAPAEGYRLAGFPDQHWREPWIIHA 193
|||||
```

```
Db 142 IYGYRVNFGVIDERLHRNREYDRYYRNINIAPAEGYRLAGFPDQHWREPWIIHA 201
QY 184 POCGSSRRITGDTCCNEETONLSTIYLRYSQSKVKRQIESDYQSEVDIYNRIDEL 240
|||||
Db 202 PQCGRSSRRITGDTCCNEETQNSTIYLRYSQSKVKRQIESDYQSEVDIYNRIDEL 258
|||||

RESULT 6
AAU00507
ID AAU00507 standard; Protein; 380 AA.
XX
XX AAU00507;
XX
XX 29-AUG-2001 (first entry)
XX
XX E. coli heat-labile enterotoxin (LT) mutant Ltdell10/112.
XX
XX Heat-labile enterotoxin; LT; LTS63Y; Ltdell10/112; mutant;
XX detoxified and immunologically active protein; ADP-ribosylation; Gs;
XX endotoxin; diarrhoea; mutein.
XX
XX Escherichia coli strain K88ac.
XX Synthetic.
XX
XX Key Location/Qualifiers
FT Misc_feature 7
FT /note= "Important residue for enzymatic activity"
FT Misc_feature 44
FT /note= "Important residue for enzymatic activity"
FT Region 58..72
FT /note= "Forms the NAD-binding site"
FT Misc_feature 61
FT /note= "Important residue for enzymatic activity"
FT Misc-difference 257
FT /note= "Encoded by TG"
XX
XX WO200119998-Al.
XX
XX 22-MAR-2001.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX 15-SEP-1999; 99WO-KR00555.
XX
XX (MOGA-) MOGAM BIOTECHNOLOGY RES INST.
XX
XX Park EJ, Kim JS, Chang J, Yum J, Chung S;
XX
XX WPI; 2001-281524/29.
XX
XX N-PSDB; AAS01506.
XX
XX New detoxified mutants of Escherichia coli heat-labile enterotoxin
XX useful as vaccine for preventing and treating diarrhoea, and as adjuvant
XX for antibody production
XX
XX Claim 6; Page 42-44; 48pp; English.
XX
XX The present sequence represents Escherichia coli heat-labile
XX enterotoxin (LT) mutant Ltdell10/112. LTS63Y (AAU00506) and Ltdell10/112
XX are two novel detoxified and immunologically active proteins (LT
XX mutants) derived by site-directed mutagenesis of the A1 subunit of wild
XX type LT. The substitution of Ser to Tyr at position 63 in LTS63Y blocks
XX NAD-binding. Deletion of Glu residues at positions 110 and 112 in
XX Ltdell10/112 eliminate the enzymatic activity of LT. The A1 subunit of
XX wild type LT catalyses ADP-ribosylation of Gs, a GTP-binding protein that
XX regulates cAMP levels. The resulting increase in cAMP is the cause of
XX diarrhoea in humans and animals e.g. pigs. The mucosal immunogenicities
XX of mutant heat-labile endotoxins LTS63Y and Ltdell10/112 were tested.
XX Groups of mice were immunised with LTS623Y or Ltdell10/112. The control
XX faecal antibody titres to LT were determined. The results showed that
XX mice immunised with LTS63Y or Ltdell10/112 contained high and
XX comparable level of anti-LT antibodies in sera and faecal extracts
```

CC compared with those immunised with wild-type LT. The LT mutants are  
CC useful as a vaccine for preventing and treating diarrhoea and as an  
CC adjuvant for antibody production.  
XX  
SQ Sequence 380 AA;

Query Match 53.3%; Score 128; DB 22; Length 380;  
Best Local Similarity 100.0%; Pred. No. 3.4e-123;  
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 113 VSALGGIPYSOIYGWYRVNFGVIDERLHNRREYDRYRNLIAPADGVRLAGFPDQ 172  
|||||  
Db 129 vsalggipysqiwyrvnfgviderlhnreaydrnlnlapedgrylgfppdhq 188  
QY 173 AMREPIWHAPCGGNSRTITGTCNEETQNLSTIYLRYSQVKRQIFSDYQSEVDI 232  
|||||  
Db 189 awreepwhapcggnssrtitgdcneetqnlstilyreyqskvkrqifsdygsevdI 248  
QY 233 YNRIRDEL 240  
|||||  
Db 249 ynrirdel 256

RESULT 7  
AAY96647  
ID AAY96647 standard; Protein; 259 AA.  
XX  
AC AAY96647;  
DT 26-SEP-2000 (first entry)  
XX  
DE Synthetic E. coli LT-A K63 mutant.  
XX  
KW Heat-labile toxin; LT-A; LT-B; mutein; transgenic plant; vaccine; oral;  
KW adjuvant; anti-bacterial; S63K.  
XX  
OS Escherichia coli.  
OS Synthetic.  
XX  
FH Key  
FT Peptide 1..19  
FT /label= signal\_peptide  
FT /note= "20"  
FT Protein 20..259  
FT /label= mature\_protein  
FT Misc-difference 82  
FT /label= S63K  
FT /note= "wild-type serine is replaced by lysine"  
XX  
PN WO200037609-A2.  
XX  
XX 29-JUN-2000.  
XX  
XX 22-DEC-1999; 99WO-US30747.  
XX  
XX 22-DEC-1998; 98US-0113507.  
XX  
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.  
PA (MASO/) MASON H S.  
PA (ARNT/) ARNTZEN C J.  
XX  
PI Mason HS, Arntzen CJ;  
XX  
XX WPI; 2000-442653/38.  
DR N-PSDB; AAA51147.  
XX  
XX New polynucleotides encoding LT-A or CT-A polypeptides for the  
PT transformation of plant cells, useful in immunogenic compositions to  
PT elicit immune responses in animals  
XX  
PS Example 2; Page -: 103pp; English.  
XX

CC This is mutant S63K Escherichia coli heat-labile toxin (LT) A subunit  
CC (LT-A). The wild-type serine was replaced with lysine at residue 63 of  
CC the mature protein, which was caused by a codon change of TCC to AAG in  
CC the coding sequence. The sequence contains plant-preferred codons and  
CC eliminates sequence motifs associated with spurious mRNA processing. A  
CC single codon insertion (CTG encoding valine) was made to accommodate the  
CC creation of a NcoI restriction site around the initiator methionine  
CC codon. Novel polynucleotides encode a mutant LT-A polypeptide or a  
CC mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A) polypeptide,  
CC which have reduced enzyme activity as compared to the wild-type LT-A or  
CC CT-A polypeptide and where at least one of the codons is altered to a  
CC plant preferred codon. The polynucleotide further comprises a nucleic  
CC acid sequence encoding LT B subunit (LT-B) or a CT B subunit (CT-B). The  
CC polynucleotides are useful for the transformation of plant cells for the  
CC production of transgenic plants to produce edible vaccines, especially  
CC oral vaccines in transgenic plants for the prophylactic or therapeutic  
CC treatment against E. coli or V. cholerae. The mutant polypeptides are  
CC also useful as adjuvants.  
CC NB: This sequence does not appear in the specification, it was made from  
CC the wild type sequence shown in AAY96647, which appears in Figure 1.  
XX  
SQ Sequence 259 AA;

Query Match 52.1%; Score 125; DB 21; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-120;  
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 64 LSLRSALHAGQSILSGYSTYIYVIATAPNMFNVDVGVSPHPYQEVSAIGGIPYSQ 123  
|||||  
Db 83 lsrsahlagqslsgystyiyviatapnmfnvndvlgvysphpyeqevsaiggiysq 142  
QY 124 IYGMVYRVNFGVIDERLHNRREYDRYRNLIAPADGVRLAGFPDQAWREPEWTHHA 183  
|||||  
Db 143 iygyrvnfgviderlhnreaydrnlnlapedgrylgfppdhqawreepwihha 202  
QY 184 PQCGG 188  
|||||  
Db 203 pqgcg 207

RESULT 8  
AAY96650  
ID AAY96650 standard; Protein; 259 AA.  
XX  
AC AAY96650;  
XX  
DT 26-SEP-2000 (first entry)  
XX  
DE Plant-optimized E. coli LT-A A63K+R192G mutant.  
XX  
KW Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;  
KW adjuvant; anti-bacterial; A63K+R192G; double mutein.  
XX  
OS Escherichia coli.  
OS Synthetic.  
XX  
FH Key  
FT Peptide 1..19  
FT /label= signal\_peptide  
FT Protein 20..259  
FT /label= mature\_protein  
FT Misc-difference 82  
FT /label= S63K  
FT /note= "the wild type serine is replaced by lysine  
FT at position 63 of the mature sequence"  
FT Misc-difference 211  
FT /label= R192G  
FT /note= "the wild type arginine is replaced by glycine  
FT at position 192 of the mature sequence"  
XX  
PN WO200037609-A2.  
XX

```
PD 29-JUN-2000.
PF 22-DEC-1999; 99WO-US30747.
XX 22-DEC-1998; 98US-0113507.
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX Mason HS, Arntzen CJ;
PI WPI; 2000-442653/38.
XX New polynucleotides encoding LT-A or CT-A polypeptides for the
DR transformation of plant cells, useful in immunogenic compositions to
XX elicit immune responses in animals
XX Example 5; Page -: 103pp; English.
XX This is Escherichia coli heat-labile toxin (LT) A subunit (LT-A) double
CC mutant S3kR192G. The wild-type serine was replaced with lysine at
CC residue 63 and wild-type arginine was replaced with glycine at residue
CC 192 of the mature protein. The coding sequence contains plant-preferred
CC codons and eliminates sequence motifs associated with spurious mRNA
CC processing. A single codon insertion (GTG encoding valine) was made to
CC accommodate the creation of a NcoI restriction site around the initiator
CC methionine codon. Novel polynucleotides encode a mutant LT-A polypeptide
CC or a mutant Vibrio cholerae cholera toxin (CT) A subunit (CT-A)
CC polypeptide, which have reduced enzyme activity as compared to the
CC wild-type LT-A or CT-A polypeptide and where at least one of the codons
CC is altered to a plant preferred codon. The polynucleotide further
CC comprises a nucleic acid sequence encoding LT B subunit (LT-B) or a CT B
CC subunit (CT-B). The polynucleotides are useful for the transformation of
CC plant cells for the production of transgenic plants to produce edible
CC vaccines, especially oral vaccines in transgenic plants for the
CC prophylactic or therapeutic treatment against E. coli or V. cholerae.
CC The mutant polypeptides are also useful as adjuvants.
CC Note: This sequence does not appear in the specification. It was
CC constructed from the wild type LT-A shown in AAY96646 which is given
CC in Figure 1 of the specification.
XX Sequence 259 AA;
SQ
Query Match 52.1%; Score 125; DB 21; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.9e-120;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 64 LSLRSALGAGSLGSGYSTYIYVIATAPNMFNVDVGVSPHPYEQEVSALGGIPYSQ 123
DB 83 lsrsahlagsgslsgystyiyviatapnmfnvndvlgvsphpqeysalggipysq 142
QY 124 IYGYWYVNFVIDERLHNRREYDRYRNINLAPAEEDGYRLAGFPDPDQAWREEPWIHHA 183
DB 143 iygywrvnfviderlhnreaydrdryrnlapiapedgyrlagfpdpdqawreepwihha 202
QY 184 PQCGG 188
DB 203 pqgcg 207
RESULT 9
AAY96649
ID AAY96649 standard; Protein: 259 AA.
XX AAY96649;
XX AC
XX 26-SEP-2000 (first entry)
XX Plant-optimized E. coli LT-A R72 mutant.
DE Heat-labile toxin; LT-A; LT-B; mutant; transgenic plant; vaccine; oral;
KW
```

```
KW adjuvant; "anti-bacterial; A72R.
XX Escherichia coli.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Peptide 1..19
FT Protein /label= signal_peptide
FT /label= mature_protein
FT Misc-difference 91
FT /label= A72R
FT /note= "Wild type alanine is replaced by arginine"
XX
XX WO200037609-A2.
XX 29-JUN-2000.
XX 22-DEC-1999; 99WO-US30747.
XX 22-DEC-1998; 98US-0113507.
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA (MASO/) MASON H S.
PA (ARNT/) ARNTZEN C J.
XX Mason HS, Arntzen CJ;
XX WPI; 2000-442653/38.
XX N-PSDB; AAA51545.
XX New polynucleotides encoding LT-A or CT-A polypeptides for the
PT transformation of plant cells, useful in immunogenic compositions to
PT elicit immune responses in animals
XX Example 4; Page -: 103pp; English.
XX This is mutant A72R Escherichia coli heat-labile toxin (LT) A subunit
CC (LT-A). The wild-type alanine was replaced with arginine at residue 72
CC of the mature protein, which was caused by a nucleotide change of GC to
CC AG at position 273-274 in the coding sequence. The sequence contains
CC plant-preferred codons and eliminates sequence motifs associated with
CC spurious mRNA processing. A single codon insertion (GTG encoding valine)
CC was made to accommodate the creation of a NcoI restriction site around
CC the initiator methionine codon. Novel polynucleotides encode a mutant
CC LT-A polypeptide or a mutant Vibrio cholerae cholera toxin (CT) A
CC subunit (CT-A) polypeptide, which have reduced enzyme activity as
CC compared to the wild-type LT-A or CT-A polypeptide and where at least one
CC of the codons is altered to a plant preferred codon. The polynucleotide
CC further comprises a nucleic acid sequence encoding LT B subunit (LT-B) or
CC a CT B subunit (CT-B). The polynucleotides are useful for the
CC transformation of plant cells for the production of transgenic plants to
CC produce edible vaccines, especially oral vaccines in transgenic plants
CC for the prophylactic or therapeutic treatment against E. coli or V.
CC cholerae. The mutant polypeptides are also useful as adjuvants.
CC Note: This sequence does not appear in the specification. It was
CC constructed from the wild type LT-A shown in AAY96646 which is given
CC in Figure 1 of the specification.
XX Sequence 259 AA;
SQ
```

Query Match 48.3%; Score 116; DB 21; Length 259;  
Best Local Similarity 100.0%; Pred. No. 5.3e-111;  
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 GOSILSGYSTYIYVIATAPNMFNVDVGVSPHPYEQEVSALGGIPYSQIYGYWYVNF 132  
DB 92 gdsilsgystyiyviatapnmfnvndvlgvsphpqeysalggipysqiygywrvnf 151  
QY 133 GVIDERLHNRREYDRYRNINLAPAEEDGYRLAGFPDPDQAWREEPWIHHA PQCGG 188  
DB 152 gviderlhnreaydrdryrnlapiapedgyrlagfpdpdqawreepwlhnapqgcg 207





CC E. coli in mammals (especially humans).

XX Sequence 240 AA;

Query Match 46.2%; Score 111; DB 19; Length 240;

Best Local Similarity 100.0%; Pred. No. 6.9e-106; Mismatches 0; Indels 0; Gaps 0;

Matches 111; Conservative 0;

Qy 1 NGDRLYRADSRPPDEIKRSGGLMPRGHNEYFDRTQNMNLYDHARGTQTGFVRYDDGYV 60

|||||

Db 1 ngdrllyradsrppdeikrsgglmprghneyfdrtqnmnlydhargtqtgfvryddgyv 60

|||||

Qy 61 STSLSLRAHLAGSILSGYSTYIYVIATAPNMFNVNDVLGVYSPHPYEQ 111

|||||

Db 61 stslslrsahlagsilsgystyiyviatapnmfnvndvlgvyspphyeq 111

RESULT 12

AAW65075

ID AAW65075 standard; protein; 240 AA.

XX AC AAW65075;

XX DT 11-SEP-1998 (first entry)

XX DE E. coli LT-A mutant A72R protein fragment.

XX KW Heat labile toxin subunit A; LT-A; mutant; immunogen; detoxification;

XX KW carrier; adjuvant; prevention; treatment; disease; diarrhoea; vaccine;

XX KW infection; enterotoxic.

XX OS Escherichia coli.

XX FH Key Location/Qualifiers

FT Misc-difference 72

FT /label= A72R

FT /note= "Wild type Ala residue is replaced by Arg"

FT

PN WC9818928-A1.

XX PD 07-MAY-1998.

XX PF 30-OCT-1997; 97WO-IB01440.

XX PR 31-OCT-1996; 96GB-0022660.

XX PA (CHIR-) CHIRON SPA.

XX PI Giullani MM, Pizza M, Rappuoli R;

XX DR WPI; 1998-272223/24.

XX PT Mutated Escherichia coli heat labile toxin subunit A - is

XX PT immunogenic and detoxified relative to wild-type, useful e.g. in

XX PT vaccines against E. coli enterotoxigenic strains and as an adjuvant

XX PS Claim 3; Page 7; 67pp; English.

XX CC This sequence represents a fragment of a mutant Escherichia coli heat

XX CC labile toxin subunit A (LT-A) where the wild type Ala residue at

XX CC position 72 is replaced by an Arg residue resulting in a toxin which

XX CC retains its immunogenicity but is detoxified. Detoxification is defined

XX CC in the specification as a reduction in toxicity relative to wild-type

XX CC toxin, such that any residual toxicity is sufficiently low to allow use

XX CC as an effective immunogenic composition in humans without significant

XX CC side effects. The protein can be combined with an acceptable carrier in

XX CC immunogenic compositions, optionally comprising an adjuvant and/or a

XX CC second immunogenic antigen. Such compositions can be administered to

XX CC prevent/treat disease in a subject e.g. traveller's diarrhoea in humans.

XX CC The protein or compositions are especially administered as vaccines

XX CC useful to prevent or treat infections by enterotoxigenic strain of

XX CC E. coli in mammals (especially humans).

CC NOTE: This sequence does not appear in the specification but has

XX been constructed from the wild-type sequence represented in AAW65074.

SQ Sequence 240 AA;

Query Match 41.2%; Score 99; DB 19; Length 240;

Best Local Similarity 100.0%; Pred. No. 1.6e-93;

Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 NREVRDYVNLNAPAEADGYRLAGPPDHQAWPEEPIHHAQCCNSRTITGDFCNE 201

|||||

Db 142 nreyrdyynlnapaeadgyrlagppdhqawreepwhhahpqqnssrtitgdtcne 201

|||||

Qy 202 ETQNLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIDEL 240

|||||

Db 202 etqnlstiylyreyqskvkrqifsdysqvsevdinyrirdel 240

RESULT 13

AAAR44016

ID AAR44016 standard; Protein; 236 AA.

XX AC AAR44016;

XX DT 08-DEC-1993 (first entry)

XX DE "Lys-63" E.coli heat labile toxin subunit A.

XX KW enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;

XX KW protomer A; site-directed mutagenesis; reduced toxicity;

XX KW ADP-ribosyltransferase activity.

XX OS Escherichia coli.

XX FH Key Location/Qualifiers

FT Misc-difference 62

FT /note= "corresponds to position 63 in cholera toxin

FT A subunit; wild-type Ser is substituted by

FT Lys to reduce toxicity"

FT

PN WC9313202-A.

XX PD 08-JUL-1993.

XX PF 30-DEC-1992; 92WO-EP03016.

XX PR 31-DEC-1991; 91IT-OMI3513.

XX PA (BIOC-) BIOGINE SCLAVO SPA.

XX PI Domenighini M, Hol W, Pizza M, Rappuoli R;

XX DR WPI; 1993-227320/28.

XX DR N-PSDB; AA051317.

XX PT Immunogenic detoxified mutant cholera toxin and heat labile toxin

XX PT - useful as vaccines against infection by Vibrio cholerae and

XX PT enterotoxin producing Escherichia coli

XX PS Claim 3; Fig 2 and Page 46; 60pp; English.

XX CC The wild-type sequence coding for the A subunit of the heat labile

XX CC toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto

XX CC et al, J Biol. Chem., 259, 5037-5044 - see AA042768) was subjected to

XX CC site-directed mutagenesis. Certain mutations were found to reduce to

XX CC toxicity (see AAR38730-32 and AAR44016-R44023). The invention relates to

XX CC immunogenic, detoxified LT-A proteins and their use in vaccines to

XX CC protect against enterotoxigenic E.coli. The amino acid sequence of

XX CC this preferred detoxified mutagen is not printed in the

XX CC specification but has been assembled from the full-length wild-type

XX CC sequence and the description given in the text. (N.B. Amino

XX CC acid numbering is based on the cholera toxin A subunit sequence).

XX	Sequence	236 AA;
SQ		
	Query Match	35.8%; Score 86; DB 14; Length 236;
	Best Local Similarity	100.0%; Pred. No. 3.7e-80;
	Matches	86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	103 VYSPHPYEVSALGGIPYSQIYGWVNFNGVIDERLHRNREYDRYYRNLIAPAEDGY 162	
DB	99 vysphpyeqevsalggipysqiygwrvnfgvlderlhrnrerydrnyrnlniapaeagy 158	
QY	163 RLAGFPDPDQAWEPPWIHAPGCG 188	
DB	159 rlagfpdpdqawreepwihapgcg 184	
	RESULT 14	
ID	AAR44017	
AC	AAR44017 standard; Protein; 236 AA.	
DT	AAR44017;	
DE	08-DEC-1993 (first entry)	
KW	"Lys-97" E.coli heat labile toxin subunit A.	
KX	enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;	
KW	protomer A; site-directed mutagenesis; reduced toxicity;	
KX	ADP-ribosyltransferase activity.	
OS	Escherichia coli.	
FH	Key Location/Qualifiers	
FT	Misc-difference 93	
FT	/note= "corresponds to position 97 in cholera toxin	
FT	A subunit; wild-type val is substituted by	
FT	Lys to reduce toxicity"	
PN	WO9313202-A.	
PD	08-JUL-1993.	
PF	30-DEC-1992; 92WO-EP03016.	
PR	31-DEC-1991; 91IT-OMI3513.	
PA	(BIOC-) BIOGINE SCLAVO SPA.	
PI	Domenighini M, Hol W, Pizza M, Rappuoli R;	
DR	N-PSDB; AAQ51318.	
PT	Immunogenic detoxified mutant cholera toxin and heat labile toxin	
PT	- useful as vaccines against infection by Vibrio cholerae and	
PT	enterotoxin producing Escherichia coli	
PS	Claim 3; Fig 2 and Page 46; 60pp; English.	
CC	The wild-type sequence coding for the A subunit of the heat labile	
CC	toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto	
CC	et al., J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to	
CC	site-directed mutagenesis. Certain mutations were found to reduce	
CC	toxicity (see AAR38730-32 and AAR44016-R44025). The invention relates to	
CC	immunogenic, detoxified LT-A proteins and their use in vaccines to	
CC	protect against enterotoxigenic E.coli. The amino acid sequence of	
CC	this preferred detoxified mtein is not printed in the	
CC	specification but has been assembled from the full-length wild-type	
CC	sequence and the description given in the text. (N.B. Amino	
CC	acid numbering is based on the cholera toxin A subunit sequence).	
XX	Sequence	236 AA;
SQ		
	Query Match	35.8%; Score 86; DB 14; Length 236;
	Best Local Similarity	100.0%; Pred. No. 3.7e-80;
	Matches	86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	103 VYSPHPYEVSALGGIPYSQIYGWVNFNGVIDERLHRNREYDRYYRNLIAPAEDGY 162	
DB	99 vysphpyeqevsalggipysqiygwrvnfgvlderlhrnrerydrnyrnlniapaeagy 158	
QY	163 RLAGFPDPDQAWEPPWIHAPGCG 188	
DB	159 rlagfpdpdqawreepwihapgcg 184	
	RESULT 14	
ID	AAR44017	
AC	AAR44017 standard; Protein; 236 AA.	
DT	AAR44017;	
DE	08-DEC-1993 (first entry)	
KW	"Lys-97" E.coli heat labile toxin subunit A.	
KX	enterotoxigenic bacteria; vaccine; immunogenic detoxified LT-A;	
KW	protomer A; site-directed mutagenesis; reduced toxicity;	
KX	ADP-ribosyltransferase activity.	
OS	Escherichia coli.	
FH	Key Location/Qualifiers	
FT	Misc-difference 93	
FT	/note= "corresponds to position 97 in cholera toxin	
FT	A subunit; wild-type val is substituted by	
FT	Lys to reduce toxicity"	
PN	WO9313202-A.	
PD	08-JUL-1993.	
PF	30-DEC-1992; 92WO-EP03016.	
PR	31-DEC-1991; 91IT-OMI3513.	
PA	(BIOC-) BIOGINE SCLAVO SPA.	
PI	Domenighini M, Hol W, Pizza M, Rappuoli R;	
DR	N-PSDB; AAQ51318.	
PT	Immunogenic detoxified mutant cholera toxin and heat labile toxin	
PT	- useful as vaccines against infection by Vibrio cholerae and	
PT	enterotoxin producing Escherichia coli	
PS	Claim 3; Fig 2 and Page 46; 60pp; English.	
CC	The wild-type sequence coding for the A subunit of the heat labile	
CC	toxin (LT-A) of a strain of E.coli known to affect humans (Yamamoto	
CC	et al., J.Biol. Chem., 259, 5037-5044 - see AAQ42768) was subjected to	
CC	site-directed mutagenesis. Certain mutations were found to reduce	
CC	toxicity (see AAR38730-32 and AAR44016-R44025). The invention relates to	
CC	immunogenic, detoxified LT-A proteins and their use in vaccines to	
CC	protect against enterotoxigenic E.coli. The amino acid sequence of	
CC	this preferred detoxified mtein is not printed in the	
CC	specification but has been assembled from the full-length wild-type	
CC	sequence and the description given in the text. (N.B. Amino	
CC	acid numbering is based on the cholera toxin A subunit sequence).	
XX	Sequence	236 AA;
SQ		



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:33:24 ; Search time 19.04 Seconds  
(without alignments)  
1211.210 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 240  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNRIRDEL 240

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 8

Total number of hits satisfying chosen parameters: 8

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	184	76.7	258	1 QLECA	heat-labile entero
2	27	11.2	125	2 A61345	heat-labile entero
3	26	10.8	258	1 XVVCA	cholera enterotoxi
4	9	3.8	46	2 S39239	cholera enterotoxi
5	8	3.3	782	2 D84514	dynamin-like prote
6	8	3.3	905	1 P3XRA4	inner capsid prote
7	8	3.3	1130	2 T23104	hypothetical prote
8	8	3.3	1133	2 T23103	hypothetical prote

ALIGNMENTS

RESULT 1  
QLECA  
heat-labile enterotoxin A precursor - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 30-Apr-1981 #sequence\_revision 17-Oct-1997 #text\_change 18-Jun-1999  
C:Accession: I55231; A01817; A26946  
R:Yamanoto, T.; Tamura, T.; Yokota, T.  
J. Biol. Chem. 259, 5037-5044, 1984  
A:Title: Primary structure of heat-labile enterotoxin produced by Escherichia coli pathd  
A:Reference number: I55231; MUID:84185610  
A:Accession: I55231  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-258 <RES>  
A:CROSS-references: GB:K01995; NID:g148027; PIDN:AAA24685.1; PID:g148028  
R:Spicer, E.K.; Noble, J.A.

J. Biol. Chem. 257, 5716-5721, 1982  
A:Title: Escherichia coli heat-labile enterotoxin. Nucleotide sequence of the A subun  
A:Reference number: A01817; MUID:82167425  
A:Accession: A01817  
A:Molecule type: DNA  
A:Residues: 1-21,'R',23-36,'FRS',40-44,46-92,'Y',94-99,'LTIYI',105-107,111-118,'IS',1  
A:CROSS-references: EMBL:V00275; NID:941339; PIDN:CAA3532.1; PID:941340  
A:Note: the authors translated the codon TAR for residue 93 as Ser  
R:Yamanoto, T.; Gojobori, T.; Yokota, T.  
J. Bacteriol. 169, 1352-1357, 1987  
A:Title: Evolutionary origin of pathogenic determinants in enterotoxigenic Escherichl  
A:Reference number: A26946; MUID:87137303  
A:Accession: A26946  
A:Molecule type: DNA  
A:Residues: 1-21,'R',23-206,'N',208-230,'E',232-255,'D',257-258 <YAM>  
A:CROSS-references: EMBL:M15363  
C:Comment: The heat-labile enterotoxin molecule contains one A chain and five or six  
class.  
C:Genetics:  
A:Gene: eta  
C:Superfamily: heat-labile enterotoxin chain A  
C:Keywords: enterotoxin  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-258/Product: heat-labile enterotoxin chain A #status predicted <MAT>

Query Match 76.7%; Score 184; DB 1; Length 258;  
Best Local Similarity 100.0%; Pred. No. 4.2e-190;  
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTFVRYDDGVVSTSL 64  
|||||  
DB 23 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMNINLYDHARGTQTFVRYDDGVVSTSL 82  
QY 65 SLRSLAGAGSILSGSYTYIYVIATAPNMFNVNDVLGVSPHPYEGEVSALGGIPYSQI 124  
|||||  
DB 83 SLRSLAGAGSILSGSYTYIYVIATAPNMFNVNDVLGVSPHPYEGEVSALGGIPYSQI 142  
QY 125 YGWRVNFVGVIDERLHRNREYRDYRNLTAPAEEDGYRLAGFPDPHQWREEPWIHAP 184  
|||||  
DB 143 YGWRVNFVGVIDERLHRNREYRDYRNLTAPAEEDGYRLAGFPDPHQWREEPWIHAP 202  
QY 185 QGCG 188  
|||||  
DB 203 QGCG 206

RESULT 2

A61345  
heat-labile enterotoxin A chain precursor - Escherichia coli (fragments)  
C:Species: Escherichia coli  
C:Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 17-Mar-1999  
C:Accession: A61345  
R:Spicer, E.K.; Kavanaugh, W.M.; Dallas, W.S.; Falkow, S.; Konigsberg, W.H.; Schafer,  
Proc. Natl. Acad. Sci. U.S.A. 78, 50-54, 1981  
A:Title: Sequence homologies between a subunits of Escherichia coli and Vibrio cholera  
A:Reference number: A61345; MUID:81223767  
A:Accession: A61345  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-125 <SPI>  
A:CROSS-references: GB:K00433  
A:Note: authors translated the codon GAG for residue 27 as Asp, CAA for residue 53 as  
C:Superfamily: heat-labile enterotoxin chain A

Query Match 11.2%; Score 27; DB 2; Length 125;  
Best Local Similarity 100.0%; Pred. No. 3.5e-21;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 NEYFDRGTQMNINLYDHARGTQTFVVR 54  
|||||  
DB 45 NEYFDRGTQMNINLYDHARGTQTFVVR 71

RESULT 3  
XVCCA  
cholera enterotoxin chain A precursor VCL1457 [validated] - Vibrio cholerae  
C:Species: Vibrio cholerae  
C:Date: 06-Jul-1982 #sequence\_revision 26-Jan-1996 #text\_change 01-Sep-2000  
C:Accession: A05129; S14623; S14625; A91268; A91746; A92298; S17665; B43864; A82  
R:Mekalanos, J.J.; Swartz, D.J.; Pearson, G.D.N.; Harford, N.; Groyne, F.; de Wilde, M.  
Nature 306, 551-557, 1983  
A:Reference number: A93320; MUID:84068199  
A:Accession: A05129  
A:Molecule type: DNA  
A:Residues: 1-258 <MEK>  
R:Dams, E.; de Wolf, M.; Dierick, W.  
submitted to the EMBL Data Library, March 1991  
A:Description: Correction of the cholera toxin nucleotide sequence of the Vibrio cholera  
A:Reference number: S14623  
A:Accession: S14623  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <DA2>  
A:Cross-references: EMBL:X58785; NID:g48420; PIDN:CAA41592.1; PID:g48421  
A:Experimental source: strain 2125  
A:Accession: S14625  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <DA2>  
A:Cross-references: EMBL:X58785; NID:g48888; PIDN:CAA41590.1; PID:g48889  
A:Experimental source: strain 569B  
R:Lai, C.Y.; Cancedda, F.; Chang, D.  
FEBS Lett. 100, 85-89, 1979  
A:Title: Primary structure of cholera toxin subunit A-1. Isolation, partial sequences an  
A:Reference number: A91268; MUID:79169830  
A:Accession: A91268  
A:Molecule type: protein  
A:Residues: 37-38, 'L', 40-44, 'SE', 47-49, 'B', 51-55, 'B', 57, 'B', 59-60, 'B', 62-66, 'Z', 68-72, 'I  
R:Duffy, L.K.; Peterson, J.W.; Kurosky, A.  
FEBS Lett. 126, 187-190, 1981  
A:Title: Isolation and characterization of a precursor form of the 'A' subunit of cholera  
A:Reference number: A91286; MUID:81212799  
A:Accession: A91286  
A:Molecule type: protein  
A:Residues: 19, 'N', 21-27 <DUF>  
R:Klapper, D.G.; Finkelstein, R.A.; Capra, J.D.  
Immunochimistry 13, 605-611, 1976  
A:Title: Subunit structure and N-terminal amino acid sequence of the three chains of cho  
A:Reference number: A91746; MUID:76259136  
A:Accession: A91746  
A:Molecule type: protein  
A:Residues: 19-36, 'R', 38, 213-232 <KUA>  
R:Duffy, L.K.; Peterson, J.W.; Kurosky, A.  
J. Biol. Chem. 256, 12252-12256, 1981  
A:Title: Covalent structure of the gamma chain of the A subunit of cholera toxin.  
A:Reference number: A92298; MUID:82053094  
A:Accession: A92298  
A:Molecule type: protein  
A:Residues: 213-246, 'ID', 249-255, 'N', 257-258 <DU2>  
R:Dams, E.; de Wolf, M.; Dierick, W.  
Biochim. Biophys. Acta 1090, 139-141, 1991  
A:Title: Nucleotide sequence analysis of the CT operon of the Vibrio cholerae classical  
A:Reference number: S17665; MUID:91355224  
A:Accession: S17665  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <DAM>  
A:Cross-references: EMBL:X58785; NID:g48888; PIDN:CAA41590.1; PID:g48889  
R:Baudry, B.; Fasano, A.; Ketley, J.; Kaper, J.B.  
Infect. Immun. 60, 428-434, 1992  
A:Title: Cloning of a gene (zot) encoding a new toxin produced by Vibrio cholerae.  
A:Reference number: A43864; MUID:92112300  
A:Accession: B43864  
A>Status: preliminary

A:Molecule type: DNA  
A:Residues: 1-6 <BAU>  
A:Cross-references: GB:M83563; NID:gl55314; PIDN:AAA27583.1; PID:gl55316  
A:Note: sequence extracted from NCBI backbone (NCBIN:77488, NCBIP:77496)  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: A82197  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-258 <HEI>  
A:Cross-references: GB:AE004224; GB:AE003852; NID:g9655952; PIDN:AAF94614.1; GSPDB:G  
A:Experimental source: serogroup O1; strain Ni961; biotype El Tor  
C:Genetics:  
A:Gene: VCL1457  
A:Map position: 1  
C:Complex: the cholera enterotoxin molecule contains three kinds of chains; an alpha  
ciate noncovalently with the subunit B, an aggregate of five beta chains  
C:Function:  
A:Description: the active component of the toxin that is primarily responsible for  $\epsilon$   
.2.5) activity also activates intracellular adenyl cyclase  
C:Superfamily: heat-labile enterotoxin chain A  
C:Keywords: enterotoxin  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-212/Product: cholera enterotoxin alpha chain #status experimental <CTA>  
F:213-258/Product: cholera enterotoxin gamma chain #status experimental <CTG>  
F:217/Disulfide bonds: Interchain (to alpha chain) #status predicted  
Query Match 10.8%; Score 26; DB 1; Length 258;  
Best Local Similarity 100.0%; Pred. No. 8.1e-20;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 29 EYFDRGTQNNINLYDHARGTQTGFVR 54  
|||||  
DB 47 EYFDRGTQNNINLYDHARGTQTGFVR 72  
RESULT 4  
S39239  
cholera enterotoxin chain A2 - Vibrio cholerae (serotype 0139)  
N:Alternate names: CTA2 cholera toxin chain A2  
C:Species: Vibrio cholerae  
A:Variety: serotype 0139  
C:Date: 08-May-1995 #sequence\_revision 26-Jul-1996 #text\_change 16-Jul-1999  
C:Accession: S39239; S39240  
R:Lebens, M.; Holmgren, J.  
submitted to the EMBL Data Library, November 1993  
A:Description: Structure and arrangement of the Cholera toxin genes in vibrio Cholera  
A:Reference number: S39238  
A:Accession: S39239  
A:Molecule type: DNA  
A:Residues: 1-46 <LEB>  
A:Cross-references: EMBL:X76390; NID:g433856; PIDN:CAA53974.1; PID:g433858; EMBL:X763  
A:Experimental source: strain 4260B; serotype 0139  
C:Superfamily: heat-labile enterotoxin chain A  
Query Match 3.8%; Score 9; DB 2; Length 46;  
Best Local Similarity 100.0%; Pred. No. 0.034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 213 EYQSKVKRQ 221  
|||||  
DB 19 EYQSKVKRQ 27  
RESULT 5  
D84514  
dynamain-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: D84514  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: D84514  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-782 <SPO>  
A:Cross-references: GB:AE002093; NID:g4587686; PIDN:AAD25856.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g14120  
A:Map position: 2

Query Match 3.3%; Score 8; DB 2; Length 782;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 QNLSTIYL 211  
|  
Db 628 QNLSTIYL 635

## RESULT 6

P3XRA4

inner capsid protein VP3 - African horse sickness virus (serotype 4)

C:Species: African horse sickness virus  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 07-May-1999  
C:Accession: B44053  
R:Iwata, H.; Yamagawa, M.; Roy, P.  
Virology 191, 251-261, 1992

A:Title: Evolutionary relationships among the gnat-transmitted orbiviruses that cause African horse sickness virus.

A:Reference number: A44053; MUID:93033117

A:Accession: B44053

A:Molecule type: genomic RNA

A:Residues: 1-905 &lt;IWA&gt;

A:Cross-references: GB:M94681

C:Superfamily: bluetongue virus core protein VP3

C:Keywords: capsid protein

Query Match 3.3%; Score 8; DB 1; Length 905;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DSRPPDEI 16  
|  
Db 791 DSRPPDEI 798

## RESULT 7

T23104

hypothetical protein H19N07.2b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T23104

R:Dobson, R.

submitted to the EMBL Data Library, March 1997

A:Reference number: Z19678

A:Accession: T23104

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1130 &lt;MIL&gt;

A:Cross-references: EMBL:292835; PIDN:CAB07397.1; GSPDB:GN00023; CESP:H19N07.2b

A:Experimental source: clone H19N07

C:Genetics:

A:Gene: CESP:H19N07.2b

A:Map position: 5  
A:Introns: 71/2; 183/2; 339/1; 659/3; 699/1; 963/2; 998/3; 1080/3; 1105/1  
  
Query Match 3.3%; Score 8; DB 2; Length 1130;  
Best Local Similarity 100.0%; Pred. No. 7.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 137 ERLHRNRE 144  
|  
Db 529 ERLHRNRE 536  
  
RESULT 8  
T23103  
hypothetical protein H19N07.2a - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T23103  
R:Dobson, R.  
submitted to the EMBL Data Library, March 1997  
A:Reference number: Z19678  
A:Accession: T23103  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1133 <MIL>  
A:Cross-references: EMBL:292835; PIDN:CAB07396.1; GSPDB:GN00023; CESP:H19N07.2a  
A:Experimental source: clone H19N07  
C:Genetics:  
A:Gene: CESP:H19N07.2a  
A:Map position: 5  
A:Introns: 71/2; 183/2; 339/1; 659/3; 699/1; 963/2; 1001/3; 1083/3; 1108/1

Query Match 3.3%; Score 8; DB 2; Length 1133;  
Best Local Similarity 100.0%; Pred. No. 7.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 ERLHRNRE 144  
|  
Db 529 ERLHRNRE 536

Search completed: June 10, 2002, 17:37:01  
Job time: 217 sec

---



GenCore version 4.5  
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:36:39 ; Search time 11.04 Seconds  
(without alignments)  
841.729 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 240  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNRDEL 240

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 8

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	100.0	258	1 ELAP_ECOLI	P06717 escherichia
2	208	86.7	258	1 ELAH_ECOLI	P43530 escherichia
3	26	10.8	258	1 CHTA_VIBCH	P01555 vibrio chol
4	11	4.6	263	1 E2BA_ECOLI	P43528 escherichia
5	8	3.3	905	1 VP3_AHSV4	P32509 african hor
6	8	3.3	905	1 VP3_AHSV6	O71025 african hor

#### ALIGNMENTS

```
RESULT 1
ELAP_ECOLI
ID ELAP_ECOLI STANDARD; PRT; 258 AA.
AC P06717: P01554:
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Heat-labile enterotoxin A chain precursor (LT-A, porcine) (LTP-A).
GN ELTA OR LTPA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RC STRAIN=ISOLATE PCG86, AND ISOLATE P307;
RX MEDLINE=87137303; Pubmed=3546273;
RA Yamamoto T., Gojobori T., Yokota T.;
RT "Evolutionary origin of pathogenic determinants in enterotoxigenic
RT Escherichia coli and Vibrio cholerae O1.";
RL J. Bacteriol. 169:1352-1357(1987).
RN [2]
```

DR PDB: 1LTI; 17-AUG-96.  
DR PDB: 1LTS; 31-JAN-94.  
DR PDB: 1LTT; 31-JAN-94.  
DR PDB: 1LT3; 07-JUL-97.  
DR PDB: 1LT4; 16-JUN-97.  
DR PDB: 1LTL; 20-APR-95.  
DR InterPro: IPR001144; Enterotoxin\_A.  
DR Pfam: PF01375; Enterotoxin\_A.  
DR PRINTS: PR00771; ENTEROTOXIN\_A.  
KW Enterotoxin; Signal; 3d-structure.  
FT SIGNAL 1 18  
FT CHAIN 19 258  
FT ACT\_SITE 205 217  
FT ACT\_SITE 130 130  
FT VARIANTS 37 39  
FT CONFLICT 45 45  
FT CONFLICT 93 93  
FT CONFLICT 100 110  
FT CONFLICT 119 120  
FT CONFLICT 159 159  
FT CONFLICT 207 207  
FT STRAND 23 27  
FT HELIX 31 37  
FT TURN 38 38  
FT STRAND 39 40  
FT TURN 43 44  
FT TURN 48 49  
FT TURN 59 64  
FT TURN 75 76  
FT STRAND 77 81  
FT HELIX 84 95  
FT STRAND 100 106  
FT TURN 110 111  
FT STRAND 112 114  
FT HELIX 115 122  
FT HELIX 126 128  
FT STRAND 130 134  
FT TURN 135 135  
FT STRAND 137 138  
FT HELIX 139 141  
FT STRAND 142 149  
FT TURN 150 151  
FT STRAND 152 153  
FT STRAND 158 159  
FT TURN 161 162  
FT HELIX 165 168  
FT TURN 169 170  
FT STRAND 174 174  
FT HELIX 176 178  
FT TURN 179 179  
FT HELIX 180 182  
FT TURN 187 188  
FT HELIX 190 193  
FT TURN 195 196  
FT HELIX 197 200  
FT TURN 203 204  
FT HELIX 215 240  
FT TURN 241 241  
FT HELIX 242 244  
FT HELIX 250 253  
SQ SEQUENCE 258 AA; 2F0786442619F81F CRC64;

Query Match 100.0%; Score 240; DB 1; Length 258;  
Best Local Similarity 100.0%; Pred. No. 2.7e-249;  
Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGRDLYRADSRPDETKRSGGLMPRGHNEYFDRTQMNLNLYDHARGTGTGFRVYDDGVV 60  
|||||  
DB 19 NGRDLYRADSRPDETKRSGGLMPRGHNEYFDRTQMNLNLYDHARGTGTGFRVYDDGVV 78  
|||||  
QY 61 STSLRSALHAGQSILSGYSTYIYVIATAPNMFNVNDVLGVSPHPYEQEVSALGGIP 120

Db 79 STSLRSALHAGQSILSGYSTYIYVIATAPNMFNVNDVLGVSPHPYEQEVSALGGIP 138  
|||||  
QY 121 YSOIYGWYRVNFGVIDERLHNRREYDRYRNLNIAPAEDGYRLAGFPDQAWREPWI 180  
|||||  
Db 139 YSOIYGWYRVNFGVIDERLHNRREYDRYRNLNIAPAEDGYRLAGFPDQAWREPWI 198  
|||||  
QY 181 HHAPOGCGNSSRTITGTCNEETONLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIDEL 240  
|||||  
Db 199 HHAPOGCGNSSRTITGTCNEETONLSTIYLREYQSKVKRQIFSDYQSEVDIYNRIDEL 258  
|||||

RESULT 2  
ELAH\_ECOLI  
ID ELAH\_ECOLI STANDARD; PRT; 258 AA.  
AC P43530;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Heat-labile enterotoxin A chain precursor (LT-A, human) (LTH-A).  
GN ELTA OR LTPA OR TOXA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE H10407;  
RX MEDLINE=84185610; PubMed=6325417;  
RA Yamamoto T., Tamura T., Yokota T.;  
RT "Primary structure of heat-labile enterotoxin produced by Escherichia  
RT coli pathogenic for humans.";  
RL J. Biol. Chem. 259:5037-5044(1984).  
RN [2]  
RP REVISION TO 207.  
RC STRAIN-ISOLATE H10407;  
RX MEDLINE=87137303; PubMed=3546273;  
RA Yamamoto T., Gojobori T., Yokota T.;  
RT "Evolutionary origin of pathogenic determinants in enterotoxigenic  
RT Escherichia coli and Vibrio cholerae O1.";  
RL J. Bacteriol. 169:1352-1357(1987).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISOLATE H10407;  
RX MEDLINE=93252225; PubMed=8486242;  
RA Inoue T., Tsuji T., Koto M., Imamura S., Miyama A.;  
RT "Amino acid sequence of heat-labile enterotoxin from chicken  
RT enterotoxigenic Escherichia coli is identical to that of human strain  
RT H 10407.";  
RL FEMS Microbiol. Lett. 108:157-161(1993).  
RN [4]  
RP DISCUSSION OF SEQUENCE.  
RX MEDLINE=95349400; PubMed=7623659;  
RA Domenighini M., Pizza M., Jobling M.G., Holmes R.K., Rappuoli R.;  
RT "Identification of errors among database sequence entries and  
RT comparison of correct amino acid sequences for the heat-labile  
RT enterotoxins of Escherichia coli and Vibrio cholerae.";  
RL Mol. Microbiol. 15:1165-1167(1995).  
CC -!- FUNCTION: THE BIOLOGICAL ACTIVITY OF THE TOXIN IS PRODUCED BY  
CC THE A CHAIN, WHICH ACTIVATES INTRACELLULAR ADENYL CYCLASE.  
CC -!- SUBUNIT: HETEROHEXAMER OF ONE A CHAIN AND OF FIVE B CHAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; K01995; AAA24685.1; -;  
DR EMBL; S60731; CAB31891.1; -;  
DR HSSP; P06717; 1LTG.

Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.

[5]  
SEQUENCE FROM N.A.  
STRAIN=KNIH002;  
Shin H.J., Park Y.C., Kim Y.C.;  
"Cloning and nucleotide sequence analysis of the virulence gene  
cassette from vibrio cholerae KNH002 isolated in Korea.";  
Misalimurhag Hoij1 35:205-210(1999).  
[6]  
SEQUENCE FROM N.A.  
STRAIN=EL TOR N16961 / SEROTYPE O1;  
MEDLINE=20406833; PubMed=10952301;  
Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
Fraser C.M.;  
"DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
*cholerae*.";  
Nature 406:477-483(2000).  
[7]  
SEQUENCE OF 1-212 FROM N.A.  
STRAIN=CLASSICAL 569B / ATCC 25870 / SEROTYPE O1;  
MEDLINE=85006737; PubMed=6090390;  
Lockman H.A., Galen J.E., Kaper J.B.;  
"Vibrio cholerae enterotoxin genes: nucleotide sequence analysis of  
DNA encoding ADP-ribosyltransferase.";  
J. Bacteriol. 159:1086-1089(1984).  
[8]

RA MEDLINE=84061784; PubMed=6315707;  
RA Lockman H., Kaper J.B.;  
RT "Nucleotide sequence analysis of the A2 and B subunits of *Vibrio*  
RT cholerae enterotoxin.";  
RL J. Biol. Chem. 258:13722-13726(1983).  
RN [9]  
RN SEQUENCE OF 19-27.  
RP MEDLINE=81212799; PubMed=7238869;  
RX Duffy L.K., Peterson J.W., Kurosky A.;  
RA "Isolation and characterization of a precursor form of the 'A'  
RT subunit of cholera toxin.";  
RT subunit of cholera toxin.";  
RL FEBS Lett. 126:187-190(1981).  
RN [10]  
RN SEQUENCE OF 19-38 AND 213-232.  
RP MEDLINE=76259136; PubMed=955672;  
RX Klapper D.G., Finkelstein R.A., Capra J.D.;  
RA "Subunit structure and N-terminal amino acid sequence of the three  
RT chains of cholera enterotoxin.";  
RL Immunochimistry 13:605-611(1976).  
RN [11]  
RN SEQUENCE OF 27-72 AND 111-139.  
RX MEDLINE=79169830; PubMed=437113;  
RA Lai C.-Y., Cancedda F., Chang D.;  
RT "Primary structure of cholera toxin subunit A1: isolation, partial  
RT sequences and alignment of the BrCN fragments.";  
RL FEBS Lett. 100:85-89(1979).  
RN [12]  
RN SEQUENCE OF 213-258.  
RX MEDLINE=82053094; PubMed=7028752;  
RD Duffy L.K., Peterson J.W., Kurosky A.;  
RA "Covalent structure of the gamma chain of the A subunit of cholera  
RT toxin.";  
RL J. Biol. Chem. 256:12252-12256(1981).  
RN [13]  
RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
RP MEDLINE=95387395; PubMed=7658473;  
RX Zhang R.G., Scott D.L., Westbrook M.L., Nance S., Spangler B.D.,  
RA Shipley G.G., Westbrook E.M.;  
RA "The three-dimensional crystal structure of cholera toxin.";  
RL J. Mol. Biol. 251:563-573(1995).  
CC -!- FUNCTION: THE ALPHA/GAMMA CHAIN (A SUBUNIT) IS AN ADP-RIBOSYLATING



```
RT hemorrhagic disease as evidenced by their capsid protein sequences." ;
RL Virology 191:251-261(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=VACCINE;
RX MEDLINE=94270993; PubMed=8002793;
RA Sakamoto K., Punyahotra R., Mizukoshi N., Ueda S., Imagawa H.,
RA Sugura T., Kanada M., Fukusho A.;
RT "Rapid detection of African horsesickness virus by the reverse
RT transcriptase polymerase chain reaction (RT-PCR) using the primer
RT for segment 3 (VP3 gene).";
RL Arch. Virol. 136:87-97(1994).
CC -1- FUNCTION: THE VP3 PROTEIN IS ONE OF THE FIVE PROTEINS (WITH VP1,
CC VP4, VP6 AND VP7) WHICH FORM THE INNER CAPSID OF THE VIRUS.
CC -1- SIMILARITY: BELONGS TO THE REOVIRUSES VP3 FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M94681; AAA42538.1; ALT_SEQ.
CC EMBL; D26572; BAA05621.1; -.
CC PIR; B44053; P3XRA4.
CC HSP; P56582; 2BTU.
CC InterPro; IPR002614; Orbi_VP3.
CC Pfam; PF01700; Orbi_VP3; 1.
CC ProDom; PD004438; Orbi_VP3; 1.
CC Core protein.
CC CONFLICT 17 18 YA -> LS (IN REF. 1).
CC CONFLICT 25 25 S -> D (IN REF. 1).
CC CONFLICT 121 121 I -> V (IN REF. 1).
CC CONFLICT 195 195 I -> V (IN REF. 1).
CC CONFLICT 231 231 Q -> P (IN REF. 1).
CC CONFLICT 253 253 F -> S (IN REF. 1).
CC CONFLICT 403 408 RALSDV -> AALTI (IN REF. 1).
CC CONFLICT 446 456 DPOTGRNGW -> THKRVGLQKV (IN REF. 2).
CC CONFLICT 560 560 F -> L (IN REF. 1).
CC CONFLICT 662 662 F -> P (IN REF. 1).
CC CONFLICT 708 708 Y -> T (IN REF. 1).
CC CONFLICT 732 732 Y -> H (IN REF. 1).
CC CONFLICT 802 850 MPVYTEERDGRVMAFKYATATAYELLYNVAYSTPTPT
CC LITVNTPT -> IANQANMGQFAAIRRTLDNGWIQFGG
CC MLRNKIKFFDSRPDEILT (IN REF. 1).
CC V -> L (IN REF. 1).
CC R -> A (IN REF. 1).
CC I -> V (IN REF. 1).
CC SEQUENCE 905 AA; 103328 MW; 92D8C3C8FAB957A0 CRC64;

Query Match 3.3%; Score 8; DB 1; Length 905;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 DSRPPDEI 16
Db 791 DSRPPDEI 798
```

```
RESULT 6
VP3_AHSV6 STANDARD; PRT; 905 AA.
AC O71025;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VP3 core protein.
GN S3 OR L3.
OS African horse sickness virus 6 (AHSV-6) (African horse sickness virus
(sero-type 6)).
```

```
OC Viruses: dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=86060;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98278331; PubMed=9617769;
RA Williams C.F., Inoue T., Lucas A.-M., Zanotto P., Roy P.;
RT "The complete sequence of four major structural proteins of African
RT horse sickness virus serotype 6: evolutionary relationships within
RT and between the orbiviruses.";
RL Virus Res. 53:53-73(1998).
CC -1- FUNCTION: THE VP3 PROTEIN IS ONE OF THE FIVE PROTEINS (WITH VP1,
CC VP4, VP6 AND VP7) WHICH FORM THE INNER CAPSID OF THE VIRUS.
CC -1- SIMILARITY: BELONGS TO THE REOVIRUSES VP3 FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF021236; AAC40995.1; -.
CC HSP; P56582; 2BTU.
CC InterPro; IPR002614; Orbi_VP3.
CC Pfam; PF01700; Orbi_VP3; 1.
CC ProDom; PD004438; Orbi_VP3; 1.
CC Core protein.
CC KW SEQUENCE 905 AA; 103305 MW; 8748565C70D9B90C CRC64;

Query Match 3.3%; Score 8; DB 1; Length 905;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 DSRPPDEI 16
Db 791 DSRPPDEI 798

Search completed: June 10, 2002, 17:40:08
Job time: 209 sec
```

"

"

\_\_\_\_\_

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:36:19 ; Search time 26.46 Seconds  
(without alignments)  
1569.115 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 240  
Sequence: 1 NGRRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNIRDEL 240

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 8

Total number of hits satisfying chosen parameters: 11

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database :
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	79.6	258	2	O66280 escherichia
2	24	10.0	68	2	Q47182 escherichia
3	9	3.8	46	2	Q57372 vibrio chol
4	8	3.3	309	17	Q9HJH5 thermoplasm
5	8	3.3	330	17	Q97AR5 thermoplasm
6	8	3.3	616	12	P87708 fowlpox vir
7	8	3.3	782	10	Q9S147 arabidopsis
8	8	3.3	903	12	Q64928 african hor
9	8	3.3	1130	5	O45624 caenorhabdi
10	8	3.3	1133	5	O45623 caenorhabdi
11	8	3.3	1766	12	Q9J599 fowlpox vir

ALIGNMENTS

```
RESULT 1
O66280 PRELIMINARY; PRT; 258 AA.
AC O66280;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HEAT-LABILE ENTEROTOXIN A SUBUNIT.
GN LTH A SUBUNIT.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1032 (ENTEROTOXIGENIC);
RA Komase K.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1032 (ENTEROTOXIGENIC);
RX MEDLINE=95091056; PubMed=7998417;
RA Tamura S., Asanuma H., Tomita T., Komase K., Kawahara K., Danbara H.,
RA Hattori N., Watanabe K., Suzuki Y., Nagamine T., Aizawa C., Oya A.,
RA Kurata T.;
RT "Escherichia coli heat-labile enterotoxin B subunits supplemented with
a trace amount of the holotoxin as an adjuvant for nasal influenza
vaccine.";
RL Vaccine 12:1083-1089(1994).
DR EMBL; AB011677; BAA25725.1; -.
DR HSP; P06717; ILTG
DR InterPro; IPR001144; Enterotoxin_A.
DR InterPro; IPR000886; ER_target.
DR Pfam; PF01375; Enterotoxin_A; 1.
DR PRINTS; PR00771; ENTEROTOXINA.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN 1.
SQ SEQUENCE 258 AA; 29931 MW; 2BB15D27740EB788 CRC64;

Query Match 79.6%; Score 191; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 2.8e-186;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMINLYDHARGTQTGFVRVDDGVYSTSL 64
DB 23 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTQMINLYDHARGTQTGFVRVDDGVYSTSL 82
QY 65 SLRSHLAGQSILSGYSTYYIYVIATAPNMFNVNDVLCGVSPHPYEQEVSALGGIPYSQI 124
DB 83 SLRSHLAGQSILSGYSTYYIYVIATAPNMFNVNDVLCGVSPHPYEQEVSALGGIPYSQI 142
QY 125 YGWRVNFVGVIDERLHNRNREYRDYRNLNIAEDGYRLAGFPDPDQWREPEWIIHAP 184
DB 143 YGWRVNFVGVIDERLHNRNREYRDYRNLNIAEDGYRLAGFPDPDQWREPEWIIHAP 202
QY 185 QGCGNSSRTIT 195
DB 203 QGCGNSSRTIT 213

RESULT 2
Q47182 PRELIMINARY; PRT; 68 AA.
AC Q47182;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENTEROTOXIN A (FRAGMENT).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
RN [1]
```

RP SEQUENCE OF 64-68 FROM N.A.  
RX STRAIN-H10407;  
RC MEDLINE=83007048; PubMed=6749816;  
RA Yamamoto T., Tamura T., Yoji M., Kaji A., Yokota T., Takano T.;  
RT "Sequence analysis of the heat-labile enterotoxin subunit B gene  
originating in human enterotoxigenic *Escherichia coli*.";  
RL J. Bacteriol. 152:506-509(1982).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H10407;  
RX MEDLINE=83114628; PubMed=6759877;  
RA Yamamoto T., Tamura T., Yokota T., Takano T.;  
RT "Overlapping genes in the heat-labile enterotoxin operon originating  
from *Escherichia coli* human strain.";  
RL Mol. Gen. Genet. 188:356-359(1982).  
RN [3]  
RP SEQUENCE OF 64-68 FROM N.A.  
RC STRAIN-H10407;  
RX MEDLINE=83265593; PubMed=6348025;  
RA Yamamoto T., Yokota T.;  
RT "Sequence of heat-labile enterotoxin of *Escherichia coli* pathogenic  
for humans.";  
RL J. Bacteriol. 155:728-733(1983).  
DR EMBL: J01646; AAB02981.1; -;  
DR HSP: P06717; 1LT3  
DR InterPro: IPR001144; Enterotoxin\_A.  
DR InterPro: IPR000886; ER\_target.  
DR Pfam: PF01375; Enterotoxin\_A; 1.  
DR PRINTS: PR00771; ENTEROTOXINA.  
DR PROSITE: PS00014; ER\_TARGET; UNKNOWN\_1.  
FT NON\_TER 1 1  
SQ SEQUENCE 68 AA; 8077 MW; 673CDE4D2E884854 CRC64;

Query Match 10.0%; Score 24; DB 2; Length 68;  
Best Local Similarity 100.0%; Pred. No. 1.2e-16;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 YQSKVKRQIFSDYQSEVDIYNRIR 237  
Db 42 YQSKVKRQIFSDYQSEVDIYNRIR 65

RESULT 3  
Q57372  
ID Q57372 PRELIMINARY; PRT; 46 AA.  
AC Q57372;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CHOLERA TOXIN A2.  
GN CTXA2.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=4260B;  
RX MEDLINE=94237453; PubMed=8181723;  
RA Lebers M., Holmgren J.;  
RT "Structure and arrangement of the cholera toxin genes in *Vibrio*  
*cholerae* O139.";  
RL FEMS Microbiol. Lett. 117:197-202(1994).  
DR EMBL: X76391; CAA53975.1; -;  
DR EMBL: X76390; CAA53974.1; -;  
DR HSP: P01555; 1XTC.  
DR InterPro: IPR001144; Enterotoxin\_A.  
DR InterPro: IPR000886; ER\_target.  
DR Pfam: PF01375; Enterotoxin\_A; 1.  
DR PROSITE: PS00014; ER\_TARGET; UNKNOWN\_1.  
SQ SEQUENCE 46 AA; 5447 MW; 1B6085A02E8889D6 CRC64;

Query Match 3.8%; Score 9; DB 2; Length 46;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 213 EYQSKVKRQ 221  
Db 19 EYQSKVKRQ 27

RESULT 4  
Q9HJH5  
ID Q9HJH5 PRELIMINARY; PRT; 309 AA.  
AC Q9HJH5;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL MEMBRANE PROTEIN.  
GN TA0993  
OS Thermoplasma acidophilum.  
OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaaceae;  
OC Thermoplasma.  
OX NCBI\_TaxID=2303;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 1728;  
RX MEDLINE=20479972; PubMed=11029001;  
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;  
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma  
acidophilum";  
RL Nature 407:508-513(2000).  
DR EMBL: AL445066; CAC12122.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 309 AA; 33633 MW; C39F9882B276B380 CRC64;

Query Match 3.3%; Score 8; DB 17; Length 309;  
Best Local Similarity 100.0%; Pred. No. 9.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 YIYIVAT 90  
Db 94 YIYIVAT 101

RESULT 5  
Q97AR5  
ID Q97AR5 PRELIMINARY; PRT; 330 AA.  
AC Q97AR5;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE TVG0750122 PROTEIN.  
GN TVG0750122.  
OS Thermoplasma volcanium.  
OC Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmaaceae;  
OC Thermoplasma.  
OX NCBI\_TaxID=50339;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GSSI / DSM 4299 / JCM 9571;  
RX MEDLINE=20570466; PubMed=11121031;  
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
RT "Archaeal adaptation to higher temperatures revealed by genomic  
sequence of Thermoplasma volcanium";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
DR EMBL: AP000993; BAB59886.1; -;  
KW Complete proteome.  
SQ SEQUENCE 330 AA; 35912 MW; 22DD4CCB803B61BB CRC64;



Query Match 3.3%; Score 8; DB 17; Length 330;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 YYIXVIAT 90  
 |||||

Db 113 YYIVVIAT 120

RESULT 6

RC P87708 PRELIMINARY; PRT; 616 AA.  
 AC P87708;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ORF1 AND ORF2 GENES, ORF3 AND ORF4 GENES (FRAGMENT).  
 OS Fowlpox virus (FPV).  
 OS Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
 OC Avipoxvirus.  
 OX NCBI\_TaxID=10261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TCP-BLEN (CEVA);  
 RA Stefancsik R., Rauch T.;  
 RT "Fowlpox virus TCP-BLEN (CEVA).";  
 RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U17141; ABA40513.1;  
 DR InterPro: IPR000209; Peptidase\_S8.  
 DR PROSITE: PS00136; SUBTILASE\_ASP; UNKNOWN\_1.  
 FT NON\_TER 616  
 SQ SEQUENCE 616 AA; 68545 MW; 3177A67B8520B24A CRC64;

Query Match 3.3%; Score 8; DB 12; Length 616;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SILSGYST 82  
 |||||

Db 272 SILSGYST 279

RESULT 7

RC Q9S147 PRELIMINARY; PRT; 782 AA.  
 AC Q9S147;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DYNAMIN-LIKE PROTEIN.  
 GN AT2G14120.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; viridiplantae; Streptophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,  
 RA Salzberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 thaliana.";  
 RL Nature 402:761-768(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007197; AAD25856.1;  
 DR InterPro: IPR001401; Dynamlin.  
 DR InterPro: IPR000375; Dynamlin\_central.  
 DR InterPro: IPR003130; GED  
 DR Pfam: PF00350; dynamlin; 1.  
 DR Pfam: PF01031; dynamlin\_2; 1.  
 DR Pfam: PF02212; GED; 1.  
 DR PRINTS; PR00195; DYNAMIN.  
 DR SMART; SM00553; DYNC; 1.  
 DR SMART; SM00302; GED; 1.  
 SQ SEQUENCE 782 AA; 86862 MW; 6789A588734EC051 CRC64;

Query Match 3.3%; Score 8; DB 10; Length 782;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 QNLSTIYL 211  
 |||||

Db 628 QNLSTIYL 635

RESULT 8

RC Q64928 PRELIMINARY; PRT; 903 AA.  
 AC Q64928;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE INNER SHELL PROTEIN.  
 GN VP3.  
 OS African horse sickness virus.  
 OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.  
 OX NCBI\_TaxID=10896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Alonso J.L., Martinez J.L., Casal J.I.;  
 RT "Nucleotide sequence of the African horsesickness virus L3 gene.";  
 RL Submitted (MAY-1992) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: M94312; AAA42542.1;  
 DR HSP: P56582; 2BT.  
 DR InterPro: IPR002614; Orbi\_VP3.  
 DR Pfam: PF01700; Orbi\_VP3; 1.  
 DR PRODOM: PD004438; Orbi\_VP3; 1.  
 SQ SEQUENCE 903 AA; 102980 MW; B261DAEF08EA42AA CRC64;

Query Match 3.3%; Score 8; DB 12; Length 903;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DSRPDEI 16  
 |||||

Db 789 DSRPDEI 796

RESULT 9

RC Q45624 PRELIMINARY; PRT; 1130 AA.  
 AC Q45624;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE H19N07.2B PROTEIN.  
 GN H19N07.2B.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]

RP SEQUENCE FROM N.A.  
RA Dobson R.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99069613; PubMed-9351916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z92835; CAB07397.1; .  
DR MEROPS; C19.UPW; .  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR002083; MATH.  
DR InterPro; IPR001394; UCH-2.  
DR Pfam; PF00917; MATH; 1.  
DR Pfam; PF00442; UCH-1; 1.  
DR Pfam; PF00443; UCH-2; 1.  
DR SMART; SM00061; MATH; 1.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
DR PROSITE; PS00972; UCH\_2\_1; 1.  
DR PROSITE; PS00973; UCH\_2\_2; 1.  
DR PROSITE; PS50235; UCH\_2\_3; 1.  
SQ SEQUENCE 1130 AA; 131082 MW; FC0F498B7925A570 CRC64;

Query Match 3.3%; Score 8; DB 5; Length 1130;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 ERLHNRRE 144  
Db 529 ERLHNRRE 536  
|||||

RESULT 10  
O45623 PRELIMINARY; PRT; 1133 AA.  
AC O45623;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE H19N07.2A PROTEIN.  
DE H19N07.2A.  
GN Caenorhabditis elegans.  
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dobson R.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99069613; PubMed-9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z92835; CAB07396.1; .  
DR MEROPS; C19.UPW; .  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR002083; MATH.  
DR InterPro; IPR001394; UCH-2.  
DR Pfam; PF00917; MATH; 1.  
DR Pfam; PF00442; UCH-1; 1.  
DR Pfam; PF00443; UCH-2; 1.  
DR SMART; SM00061; MATH; 1.  
DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_1.  
DR PROSITE; PS00972; UCH\_2\_1; 1.  
DR PROSITE; PS00973; UCH\_2\_2; 1.  
DR PROSITE; PS50235; UCH\_2\_3; 1.  
SQ SEQUENCE 1133 AA; 131355 MW; 6742C643285C6025 CRC64;

Query Match 3.3%; Score 8; DB 5; Length 1133;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 ERLHNRRE 144  
Db 529 ERLHNRRE 536  
|||||

RESULT 11  
Q9J599 PRELIMINARY; PRT; 1766 AA.  
AC Q9J599;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE ORF FPV123 VARIOLA B22R GENE FAMILY PROTEIN.  
DE GN FPV123.  
OS Fowlpox virus (FPV).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OC Avipoxvirus.  
OX NCBI\_TaxID=10261;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20193820; PubMed-10729156;  
RA Afonso C.L.; Tulman E.R.; Lu Z.; Zsak L.; Kutish G.F.; Rock D.L.;  
RT "The genome of fowlpox virus."  
RL J. Virol. 74:3815-3831(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Afonso C.L.; Tulman E.R.; Lu Z.; Zsak L.; Kutish G.F.; Rock D.L.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF198100; AAF44467.1; .  
DR InterPro; IPR000209; Peptidase\_S8.  
DR PROSITE; PS00136; SUBTILASE\_ASP; UNKNOWN\_1.  
SQ SEQUENCE 1766 AA; 201146 MW; B3D87FF4A52679D9 CRC64;

Query Match 3.3%; Score 8; DB 12; Length 1766;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 SILSGYST 82  
Db 728 SILSGYST 735  
|||||

Search completed: June 10, 2002, 17:39:51  
Job time: 212 sec





OLIGO-INTERFERENCE

us-09-297-171-1-oligo.ra1

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2002, 17:33:09 ; Search time 13.05 Seconds  
(without alignments)  
449.206 Million cell updates/sec

Title: US-09-297-171-1  
Perfect score: 240  
Sequence: 1 NGDRLYRADSRPPDEIKRSG.....QIFSDYQSEVDIYNRDEL 240

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 8

Total number of hits satisfying chosen parameters: 49

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	188	78.3	240	4	US-08-823-120-3
2	86	35.8	236	4	US-08-823-120-6
3	86	35.8	254	3	US-08-486-099-114
4	86	35.8	254	3	US-08-360-107A-124
5	86	35.8	254	3	US-08-484-223B-114
6	86	35.8	254	3	US-08-919-597-114
7	86	35.8	254	3	US-08-475-668A-114
8	86	35.8	254	3	US-08-485-551A-114
9	86	35.8	254	3	US-08-471-913A-114
10	86	35.8	254	4	US-08-485-264A-114
11	86	35.8	254	4	US-08-474-349A-114
12	26	10.8	179	2	US-08-435-605A-57
13	26	10.8	192	2	US-08-435-605A-8
14	26	10.8	194	2	US-08-435-605A-6
15	26	10.8	194	2	US-08-435-605A-14
16	26	10.8	194	2	US-08-435-605A-16
17	26	10.8	194	2	US-08-435-605A-54
18	26	10.8	194	2	US-08-435-605A-55
19	26	10.8	238	4	US-08-823-120-2
20	26	10.8	240	2	US-08-435-605A-13
21	26	10.8	240	2	US-08-435-605A-15
22	26	10.8	240	2	US-08-435-605A-50
23	26	10.8	240	2	US-08-435-605A-51
24	26	10.8	240	2	US-08-435-605A-52
25	26	10.8	240	2	US-08-435-605A-53
26	26	10.8	240	2	US-08-435-605A-56
27	26	10.8	240	4	US-08-823-120-4

28	26	10.8	240	4	US-08-823-120-8	Sequence 8, Appli
29	26	10.8	258	1	US-08-449-045C-2	Sequence 2, Appli
30	26	10.8	258	2	US-08-435-605A-2	Sequence 2, Appli
31	26	10.8	258	6	5223610-2	Patent No. 5223610
32	22	9.2	240	2	US-08-435-605A-49	Sequence 49, Appli
33	20	8.3	20	1	US-08-171-299B-11	Sequence 11, Appli
34	15	6.2	15	3	US-08-296-848A-2	Sequence 2, Appli
35	12	5.0	12	2	US-08-292-968-7	Sequence 7, Appli
36	12	5.0	12	2	US-08-467-974-7	Sequence 7, Appli
37	12	5.0	12	2	US-08-467-536-7	Sequence 7, Appli
38	12	5.0	12	3	US-08-467-976-7	Sequence 7, Appli
39	12	5.0	12	4	US-09-082-514-7	Sequence 7, Appli
40	12	5.0	46	2	US-08-435-605A-10	Sequence 10, Appli
41	9	3.8	9	2	US-08-292-968-8	Sequence 8, Appli
42	9	3.8	9	2	US-08-467-974-8	Sequence 8, Appli
43	9	3.8	9	2	US-08-467-536-8	Sequence 8, Appli
44	9	3.8	9	3	US-08-467-976-8	Sequence 8, Appli
45	9	3.8	9	4	US-09-082-514-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-08-823-120-3  
; Sequence 3, Application US/08823120  
; Patent No. 6149919  
; GENERAL INFORMATION:  
; APPLICANT: Domenighini, Mario  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Pizza, Mariagrazia  
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of  
; TITLE OF INVENTION: Cholera Toxin and of the Toxin It, Their Preparation and  
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,120  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,003  
; FILING DATE: 11-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0315.001  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 240 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-823-120-3

Query Match 78.3%; Score 188; DB 4; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.8e-181;

Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NGRDLYRADSRPDEIKRSGGLMPRGHNEYFDRCTQMNILYDHARGTGTGFVRYDDGYV 60  
Db 1 NGRDLYRADSRPDEIKRSGGLMPRGHNEYFDRCTQMNILYDHARGTGTGFVRYDDGYV 60

Qy 61 STSLSLSAHLAGOSILSGYSTYIIYVIATAPNMFNVNDVLGYSPHPYQEVSAALGGIP 120  
Db 61 STSLSLSAHLAGOSILSGYSTYIIYVIATAPNMFNVNDVLGYSPHPYQEVSAALGGIP 120

Qy 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNLIAPAEDGYRLAGFPDPHQAWREEPWI 180  
Db 121 YSQIYGWYRVNFGVIDERLHRNREYRDYRNLIAPAEDGYRLAGFPDPHQAWREEPWI 180

Qy 181 HHAPQGGC 188  
Db 181 HHAPQGGC 188

RESULT 2  
US-08-823-120-6  
; Sequence 6, Application US/08823120  
; Patent No. 6149919  
; GENERAL INFORMATION:  
; APPLICANT: Domenighini, Mario  
; APPLICANT: Rappuoli, Rino  
; APPLICANT: Piazza, Mariagrazia  
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of  
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and  
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/823,120  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/256,003  
; FILING DATE: 11-NOV-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0315.001  
; TELEPHONE: (510) 601-2703  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-823-120-6

Query Match 35.8%; Score 86; DB 4; Length 236;  
Best Local Similarity 100.0%; Pred. No. 7.8e-79;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 VYSPHPYQEVSAALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLIAPAEDGY 162  
Db 99 VYSPHPYQEVSAALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLIAPAEDGY 158

Qy 163 RLAGFPDPHQAWREEPWIHHAPQGGC 188  
Db 159 RLAGFPDPHQAWREEPWIHHAPQGGC 184

RESULT 3  
US-08-486-099-114  
; Sequence 114, Application US/08486099  
; Patent No. 6013263  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS  
; TITLE OF INVENTION: B VIRUS TRANSMISSION  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,099  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-031  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-486-099-114

Query Match 35.8%; Score 86; DB 3; Length 254;  
Best Local Similarity 100.0%; Pred. No. 8.3e-79;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 103 VYSPHPYQEVSAALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLIAPAEDGY 162  
Db 117 VYSPHPYQEVSAALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLIAPAEDGY 176

Qy 163 RLAGFPDPHQAWREEPWIHHAPQGGC 188  
Db 177 RLAGFPDPHQAWREEPWIHHAPQGGC 202

RESULT 4  
US-08-360-107A-124  
; Sequence 124, Application US/08360107A  
; Patent No. 6017536

```

; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-484-223B-114

Query Match 35.8%; Score 86; DB 3; Length 254;
Best Local Similarity 100.0%; Pred. No. 8.3e-79;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps

QY 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNPGVTDRLHRNREYRDYRNLTAPAEDGY 162
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNPGVTDRLHRNREYRDYRNLTAPAEDGY 176
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 163 RLAGFPDHOAWREEPWIHAPOGCG 188
Db 177 RLAGFPDHOAWREEPWIHAPOGCG 202

RESULT 6
; US-08-919-597-114
; Sequence 114, Application US/08919597
; Patent No. 6054265
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/919,597  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/470,896  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-919-597-114

Query Match 35.8%; Score 86; DB 3; Length 254;  
Best Local Similarity 100.0%; Pred. No. 8.3e-79;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPADGY 162  
|||||  
DB 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPADGY 176  
QY 163 RLAGFPDPHQAWREPWIHAFQCG 188  
|||||  
DB 177 RLAGFPDPHQAWREPWIHAFQCG 202

RESULT 7  
US-08-475-668A-114  
Sequence 114, Application US/08475668A  
Patent No. 6060065  
GENERAL INFORMATION:  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE  
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 211  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,668A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7872-026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-475-668A-114

Query Match 35.8%; Score 86; DB 3; Length 254;  
Best Local Similarity 100.0%; Pred. No. 8.3e-79;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPADGY 162  
|||||  
DB 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPADGY 176  
QY 163 RLAGFPDPHQAWREPWIHAFQCG 188  
|||||  
DB 177 RLAGFPDPHQAWREPWIHAFQCG 202

RESULT 8  
US-08-485-551A-114  
Sequence 114, Application US/08485551A  
Patent No. 6068973  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE  
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 211  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,551A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:



```

; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-114

Query Match 35.8%; Score 86; DB 3; Length 254;
Best Local Similarity 100.0%; Pred. No. 8.3e-79;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQEVSALGGIPYQIYGYWRVNFVGVIDERLHRNREYDRYRNLIAPAEDGY 162
|
Db 117 VYSPHPYEQEVSALGGIPYQIYGYWRVNFVGVIDERLHRNREYDRYRNLIAPAEDGY 176
|
QY 163 RLAGFPDPHQAWREEPWIHPAQCGG 188
|
Db 177 RLAGFPDPHQAWREEPWIHPAQCGG 202
|
RESULT 9
US-08-471-913A-114
; Sequence 114, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/471,913A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-471-913A-114

Query Match 35.8%; Score 86; DB 3; Length 254;
Best Local Similarity 100.0%; Pred. No. 8.3e-79;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQEVSALGGIPYQIYGYWRVNFVGVIDERLHRNREYDRYRNLIAPAEDGY 162
|
Db 117 VYSPHPYEQEVSALGGIPYQIYGYWRVNFVGVIDERLHRNREYDRYRNLIAPAEDGY 176
|
QY 163 RLAGFPDPHQAWREEPWIHPAQCGG 188
|
Db 177 RLAGFPDPHQAWREEPWIHPAQCGG 202
|
RESULT 11
US-08-474-349A-114
```

```

|
Db 117 VYSPHPYEQEVSALGGIPYQIYGYWRVNFVGVIDERLHRNREYDRYRNLIAPAEDGY 176
|
QY 163 RLAGFPDPHQAWREEPWIHPAQCGG 188
|
Db 177 RLAGFPDPHQAWREEPWIHPAQCGG 202
|
RESULT 10
US-08-485-264A-114
; Sequence 114, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/485,264A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-264A-114

Query Match 35.8%; Score 86; DB 4; Length 254;
Best Local Similarity 100.0%; Pred. No. 8.3e-79;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQEVSALGGIPYQIYGYWRVNFVGVIDERLHRNREYDRYRNLIAPAEDGY 162
|
Db 117 VYSPHPYEQEVSALGGIPYQIYGYWRVNFVGVIDERLHRNREYDRYRNLIAPAEDGY 176
|
QY 163 RLAGFPDPHQAWREEPWIHPAQCGG 188
|
Db 177 RLAGFPDPHQAWREEPWIHPAQCGG 202
|
RESULT 11
US-08-474-349A-114
```





---

## SUMMARIES

```
QY 61 STSLSLSAHLAQSILSGYSTYYIVYIATAPNMFNVNDLVGYSPHPYEQEVSALGGIP 120
Db 61 STSLSLSAHLAQSILSGYSTYYIVYIATAPNMFNVNDLVGYSPHPYEQEVSALGGIP 120

QY 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 180

QY 181 HHAPQGCNSRTTGTCTCNEETONLSTIYLREYQSKVKQIFSDYOSEVDIYNRIDEL 240
Db 181 HHAPQGCNSRTTGTCTCNEETONLSTIYLREYQSKVKQIFSDYOSEVDIYNRIDEL 240

RESULT 2
US-08-256-003-3
; Sequence 3, Application US/08256003
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt, Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,003
; FILING DATE: 11-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-256-003-3

Query Match 78.3%; Score 188; DB 6; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e-185;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NGRDLRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 60
Db 1 NGRDLRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYV 60

QY 61 STSLSLSAHLAQSILSGYSTYYIVYIATAPNMFNVNDLVGYSPHPYEQEVSALGGIP 120
Db 61 STSLSLSAHLAQSILSGYSTYYIVYIATAPNMFNVNDLVGYSPHPYEQEVSALGGIP 120

QY 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 180
Db 121 YSQIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWI 180
```

```
QY 181 HHAPQSCG 188
Db 181 HHAPQCG 188

RESULT 3
PCT-US99-30747-4
; Sequence 4, Application PC/TUS9930747
; GENERAL INFORMATION:
; APPLICANT: Boyce Thompson Institute for Plant Research at Cor
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/85427
; CURRENT APPLICATION NUMBER: PCT/US99/30747
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
; PCT-US99-30747-4

Query Match 76.7%; Score 184; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.6e-181;
Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYVSTSL 64
Db 24 LYRADSRPPDEIKRSGGLMPRGHNEYFDRGTOMNINLYDHARGTQTGFVRYDDGYVSTSL 83

QY 65 SLRSAHLAQSILSGYSTYYIVYIATAPNMFNVNDLVGYSPHPYEQEVSALGGIPYSQI 124
Db 84 SLRSAHLAQSILSGYSTYYIVYIATAPNMFNVNDLVGYSPHPYEQEVSALGGIPYSQI 143

QY 125 YGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWIHHAP 184
Db 144 YGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPHQAWREEPWIHHAP 203

QY 185 QCGG 188
Db 204 QCGG 207

RESULT 4
US-09-470-124-4
; Sequence 4, Application US/09470124
; GENERAL INFORMATION:
; APPLICANT: Mason
; APPLICANT: Arntzen
; TITLE OF INVENTION: Orally Immunogenic Bacterial Enterotoxins Expressed In
; TITLE OF INVENTION: Transgenic Plants
; FILE REFERENCE: 4868/84454
; CURRENT APPLICATION NUMBER: US/09/470,124
; CURRENT FILING DATE: 1999-12-24
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:E. coli
; OTHER INFORMATION: heat-labile toxin gene mutagenized to optimize
; OTHER INFORMATION: expression in plants.
; US-09-470-124-4
```

```
Query Match          76.7%; Score 184; DB 18; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.6e-181; Indels 0; Gaps 0;
Matches 184; Conservative 0; Mismatches 0;

QY 5 LYRADSRPDEIKRSGLMPRGHNEFYDRGTQMNINLYDHARTQTGCFVRYDDGYVSTSL 64
   |||||||
Db 24 LYRADSRPDEIKRSGLMPRGHNEFYDRGTQMNINLYDHARTQTGCFVRYDDGYVSTSL 83
   |||||||

QY 65 SLRSAHLAQSILSGYSTYYIYVIATAPNMFNVNDVLGYSPHPYQEVSAALGGIPYSOI 124
   |||||||
Db 84 SLRSAHLAQSILSGYSTYYIYVIATAPNMFNVNDVLGYSPHPYQEVSAALGGIPYSOI 143
   |||||||

QY 125 YGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPDQAWREEPWIHHAP 184
   |||||||
Db 144 YGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGYRLAGFPDPDQAWREEPWIHHAP 203
   |||||

QY 185 QGCG 188
   |||||
Db 204 QGCG 207
   |||||

RESULT 5
US-09-819-917-7
; Sequence 7, Application US/09819917
; GENERAL INFORMATION:
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Fontana, Maria Rita
; APPLICANT: Giannelli, Valentina
; APPLICANT: Rappuoli, Rino
; TITLE OF INVENTION: Immunogenic Detoxified Mutants Of Cholera Toxin
; FILE REFERENCE: CHIR0312
; CURRENT APPLICATION NUMBER: US/09/819,917
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 08/981,208
; PRIOR FILING DATE: 1997-12-22
; PRIOR APPLICATION NUMBER: 9513371.6
; PRIOR FILING DATE: 1995-06-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 7
; LENGTH: 240
; TYPE: PRT
; ORGANISM: E. coli
US-09-819-917-7

Query Match          45.0%; Score 108; DB 22; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.2e-103; Indels 0; Gaps 0;
Matches 108; Conservative 0; Mismatches 0;

QY 81 STYIIVVIATAPNMFNVNDVLGYSPHPYQEVSAALGGIPYSOIYGWYRVNFGVIDERLH 140
   |||||||
Db 81 STYIIVVIATAPNMFNVNDVLGYSPHPYQEVSAALGGIPYSOIYGWYRVNFGVIDERLH 140
   |||||||

QY 141 RNREYDRYRNINLPAEDGYRLAGFPDPDQAWREEPWIHHAPQCG 188
   |||||||
Db 141 RNREYDRYRNINLPAEDGYRLAGFPDPDQAWREEPWIHHAPQCG 188
   |||||||

RESULT 6
US-08-256-003-6
; Sequence 6, Application US/08256003
; GENERAL INFORMATION:
; APPLICANT: Domenighini, Mario
; APPLICANT: Rappuoli, Rino
; APPLICANT: Pizza, Mariagrazia
; TITLE OF INVENTION: Immunogenic Detoxified Mutants of
; TITLE OF INVENTION: Cholera Toxin and of the Toxin Lt. Their Preparation and
; TITLE OF INVENTION: Their Use for the Preparation of Vaccines
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
```

```
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/256,003
; FILING DATE: 11-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0315.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-256-003-6

Query Match          35.8%; Score 86; DB 6; Length 236;
Best Local Similarity 100.0%; Pred. No. 5e-80; Indels 0; Gaps 0;
Matches 86; Conservative 0; Mismatches 0;

QY 103 VYSPHPYQEVSAALGGIPYSOIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGY 162
   |||||||
Db 99 VYSPHPYQEVSAALGGIPYSOIYGWYRVNFGVIDERLHRNREYDRYRNINLPAEDGY 158
   |||||||

QY 163 RLAGFPDPDQAWREEPWIHHAPQCG 188
   |||||||
Db 159 RLAGFPDPDQAWREEPWIHHAPQCG 184
   |||||||

RESULT 7
US-09-044-696-2
; Sequence 2, Application US/09044696
; GENERAL INFORMATION:
; APPLICANT: BARCHFELD, GAIL
; APPLICANT: DEL GIUDICE, GIUSEPPE
; APPLICANT: RAPPUOLI, RINO
; TITLE OF INVENTION: DETOXIFIED MUTANTS OF BACTERIAL
; TITLE OF INVENTION: ADP-RIBOSYLATING TOXINS AS PARENTERAL ADJUVANTS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION, INTELLECTUAL PROPERTY -
; STREET: P.O. BOX 8097
; CITY: EMERYVILLE
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/044,696
; FILING DATE: 18-MAR-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/041,227
; FILING DATE: 21-MAR-1997
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 1393.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 655-8730
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-044-696-2

Query Match 35.8%; Score 86; DB 14; Length 236;
Best Local Similarity 100.0%; Pred. No. 5e-80;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHNRNREYDRYRNLTAPADGY 162
Db 99 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHNRNREYDRYRNLTAPADGY 158
QY 163 RLAGFPDPHQAWREPWIHHAPOCG 188
Db 159 RLAGFPDPHQAWREPWIHHAPOCG 184

RESULT 8
US-08-360-107-124
; Sequence 124, Application US/08360107
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 137
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/360,107
; FILING DATE: 20-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 124:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; STRANDEDNESS:
; TOPOLOGY: unknown
```

```
; MOLECULE TYPE: protein
US-08-360-107-124

Query Match 35.8%; Score 86; DB 7; Length 254;
Best Local Similarity 100.0%; Pred. No. 5.4e-80;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHNRNREYDRYRNLTAPADGY 162
Db 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHNRNREYDRYRNLTAPADGY 176
QY 163 RLAGFPDPHQAWREPWIHHAPOCG 188
Db 177 RLAGFPDPHQAWREPWIHHAPOCG 202

RESULT 9
US-08-470-896-114
; Sequence 114, Application US/08470896
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,896
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-470-896-114

Query Match 35.8%; Score 86; DB 8; Length 254;
Best Local Similarity 100.0%; Pred. No. 5.4e-80;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHNRNREYDRYRNLTAPADGY 162
Db 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHNRNREYDRYRNLTAPADGY 176
```



QY 163 RLAGFPDHOAWREEPWIHAPQCG 188  
|||||  
Db 177 RLAGFPDHOAWREEPWIHAPQCG 202

## RESULT 10

US-08-471-913-114  
; Sequence 114, Application US/08471913  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE  
; FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS  
; TITLE OF INVENTION: TRANSMISSION  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US/08/471,913  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-030  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-471-913-114

Query Match 35.8%; Score 86; DB 8: Length 254;  
Best Local Similarity 100.0%; Pred. No. 5.4e-80;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQVVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPAEDGY 162  
|||||  
Db 117 VYSPHPYEQVVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPAEDGY 176

QY 163 RLAGFPDHOAWREEPWIHAPQCG 188  
|||||  
Db 177 RLAGFPDHOAWREEPWIHAPQCG 202

## RESULT 11

US-08-475-668-114  
; Sequence 114, Application US/08475668  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.

; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA  
; TITLE OF INVENTION: VIRUS TRANSMISSION  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: US/08/475,668  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-026  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 254 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-475-668-114

Query Match 35.8%; Score 86; DB 8: Length 254;  
Best Local Similarity 100.0%; Pred. No. 5.4e-80;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQVVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPAEDGY 162  
|||||  
Db 117 VYSPHPYEQVVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDYRNLTAPAEDGY 176

QY 163 RLAGFPDHOAWREEPWIHAPQCG 188  
|||||  
Db 177 RLAGFPDHOAWREEPWIHAPQCG 202

## RESULT 12

US-08-484-223-114  
; Sequence 114, Application US/08484223  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TITLE OF INVENTION: TRANSMISSION  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,223  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-029  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-223A-114

Query Match 35.8%; Score 86; DB 8; Length 254;  
Best Local Similarity 100.0%; Pred. No. 5.4e-80;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDRYRNLTAPAEQGY 162  
|||||  
Db 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDRYRNLTAPAEQGY 176  
|||||  
QY 163 RLAGFPDQHWREPWIIHAPQCG 188  
|||||  
Db 177 RLAGFPDQHWREPWIIHAPQCG 202

RESULT 13  
US-08-484-223A-114  
; Sequence 114, Application US/08484223A  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TITLE OF INVENTION: TRANSMISSION  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,223A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-029  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-223A-114

Query Match 35.8%; Score 86; DB 8; Length 254;  
Best Local Similarity 100.0%; Pred. No. 5.4e-80;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDRYRNLTAPAEQGY 162  
|||||  
Db 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYRDRYRNLTAPAEQGY 176  
|||||  
QY 163 RLAGFPDQHWREPWIIHAPQCG 188  
|||||  
Db 177 RLAGFPDQHWREPWIIHAPQCG 202

RESULT 14  
US-08-485-546-114  
; Sequence 114, Application US/08485546  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Dani P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Petteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE  
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS  
; TITLE OF INVENTION: TRANSMISSION  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,546  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-028  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-485-546-114

Query Match 35.8%; Score 86; DB 8; Length 254;  
Best Local Similarity 100.0%; Pred. No. 5.4e-80;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYDRYRNLIAPAEDGY 162  
Db 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYDRYRNLIAPAEDGY 176  
Qy 163 RLAGFPDPHQAWREEPWIHHAPOCGG 188  
Db 177 RLAGFPDPHQAWREEPWIHHAPOCGG 202

RESULT 15  
US-08-485-546A-114  
Sequence 114, Application US/08485546A  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE  
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS  
NUMBER OF SEQUENCES: 214  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,546A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-028  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 254 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-485-546A-114

Query Match 35.8%; Score 86; DB 8; Length 254;  
Best Local Similarity 100.0%; Pred. No. 5.4e-80;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 103 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYDRYRNLIAPAEDGY 162  
Db 117 VYSPHPYEQEVSALGGIPYSQIYGWYRVNFGVIDERLHRNREYDRYRNLIAPAEDGY 176  
Qy 163 RLAGFPDPHQAWREEPWIHHAPOCGG 188  
Db 177 RLAGFPDPHQAWREEPWIHHAPOCGG 202

Search completed: June 10, 2002, 17:38:55  
Job time: 216 sec



```

RESULT      1
US-09-528-682-1
; Sequence 1, Application US/09528682
; GENERAL INFORMATION:
; APPLICANT: Pizza, Mariagrazia
; APPLICANT: Giuliani, Marzia M
; APPLICANT: Reppulli, Rino
; TITLE OF INVENTION: IMMUNOGENIC DETOXIFIED MUTANT E. COLI LT-A-TOXIN
; FILE REFERENCE: 2302-0342.10
; CURRENT APPLICATION NUMBER: US/09/528,682
; CURRENT FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: PCT/IB97/01440
; PRIOR FILING DATE: 1997-10-30
; PRIOR APPLICATION NUMBER: 09/297,171
; PRIOR FILING DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0

```





